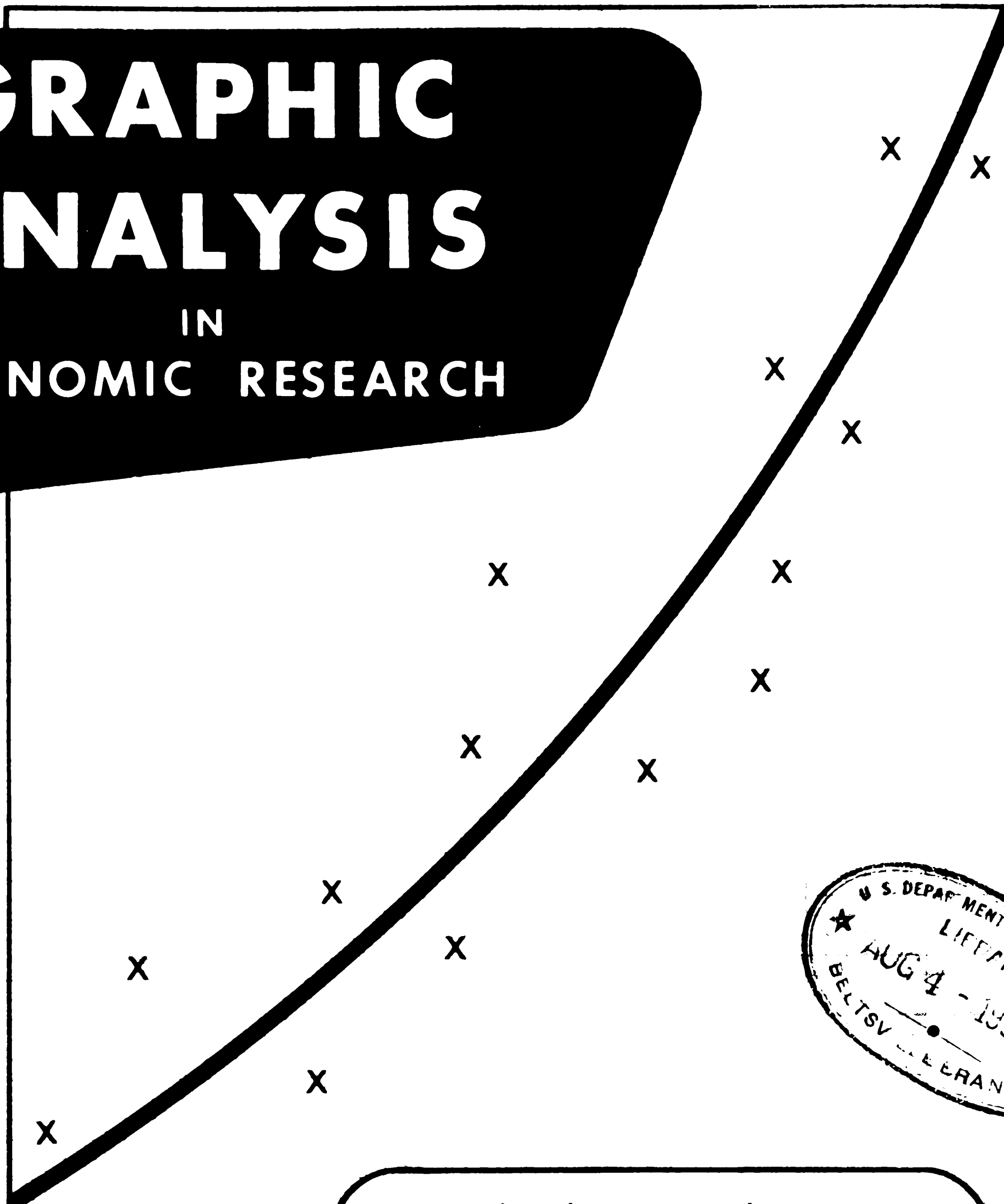
GRAPHIC ANALYSIS IN ECONOMIC RESEARCH



by Frederick V. Waugh, Director
Division of Agricultural Economics

AGRICULTURAL HANDBOOK No. 84
UNITED STATES DEPARTMENT OF AGRICULTURE
Agricultural Marketing Service
Washington, D. C.
June 1955

CONTENTS

For sale by the Superintendent of Documents, U. S. Government Printing Office, Washington 25, D. C. -Price 40 cents

GRAPHIC ANALYSIS IN ECONOMIC RESEARCH

By Frederick V. Waugh, Director,
Division of Agricultural Economics,
Agricultural Marketing Service 1/

INTRODUCTION

Graphic analysis can be a powerful tool in economic research. Economic statisticians have allowed this tool to become rusty in recent years. They have been fascinated with new tools--with new mathematical theories and with wonderful electronic computing machines. The new tools are good, too. We do not suggest that any analyst throw away his mathematical textbooks or discard his calculating machine. On the other hand, we do suggest that all economists and statisticians should be familiar with graphic analysis, and should use it along with the other tools. The purpose of this handbook is to review some of the graphic tools that can be useful. It is not a complete catalog. Rather, it is a collection of examples covering a variety of analyses.

In the two decades from 1920 to 1940, graphic analysis was very popular, especially among agricultural economists. Some analysts used graphics --and graphics alone--to measure demand curves, supply curves, cost curves, and input-output relationships. But sometimes they misused graphic methods and got spurious results. Malenbaum and Black, 2/ in 1937, warned against the indiscriminate use of graphic analysis. Without doubt, there was merit in these warnings; some practitioners of the graphic arts had done slipshod, unreliable work.

Perhaps partly because of its misuse, graphic analysis seems to be under an eclipse today. Many of the recent books on econometrics--that is, books that deal with the use of economic theory, mathematics, and statistics to measure the influence of economic forces--are filled with high-powered mathematics and a few results worked out on calculating machines, but use little or no graphic analysis. A notable exception is Tinbergen. 3/

In the Division of Agricultural Economics, we never have abandoned graphic analysis, although we, too, have been guilty of some neglect. Many of our research bulletins would be more readable, and more accurate, if we had used more diagrams and probably somewhat less detailed results ground out of the calculating machines. Often there can be a fictitious accuracy in results computed on the machine to six or eight "significant figures." Even the modern electronic computer cannot give results that are any more accurate than the numbers we put into it. The economist often must use data that are reliable only to two or three significant figures.

1/ Several members of the staffs of the Agricultural Marketing Service and the Agricultural Research Service provided material used in this report. The author thanks especially Richard J. Foote and Hyman Weingarten for preparing the material for publication.

2/ Malenbaum, I. W., and Black, J. D. The Short-Cut Graphic Method. Quart. Jour. Econ. 52:66-112, illus. 1937.

3/ Tinbergen, J. Econometrics. Transl. from the Dutch by H. Ryken Van Olst. 258 pp., illus. Philadelphia, Pa. 1951.

Graphic analysis is easy and flexible. Some have said that it is too easy, and too flexible. It can go wild if it is not combined with sound economic theory and with an understanding of probable errors. But so can any method of analysis. Especially when the economist mechanically extrapolates a trend into the future, he is jumping blindfolded into the wild blue yonder. Unless he understands the forces that made the trend in the past and are likely to shape it in the future, the fanciest mathematical function computed to six decimals is as likely to lead him astray as is the easy, flexible graphic method. Ease and flexibility are desirable features of any method. Would anyone seriously argue that difficulty and inflexibility are virtues?

Two statisticians in the former Bureau of Agricultural Economics were closely identified with the development of graphic analysis. One of them, Louis Bean, used graphics almost exclusively. The other, Mordecai Ezekiel, used graphic methods to supplement machine computation. 4/ He often computed certain linear functions on the machine, and then used graphic methods to secure a better fit by making use of curvilinear relations. This is a sound procedure if used with judgment and in moderation.

Graphic methods can and should be used to see whether mathematical results appear to fit the data. This approach has been followed in certain research studies issued in the last few years by the Bureau of Agricultural Economics and its successor, the Agricultural Marketing Service. 5/ Too often, nowadays, econometricians simply assume that economic relationships are linear in arithmetic or logarithmic terms and never bother to test whether the assumption is correct. Actually, few economic relationships are likely to be strictly linear. Graphics can often provide more meaningful relations than those given directly by machine computation. In spite of some of the publicity, not even the newest calculating machine can think. The economist must do it. Graphic analysis can help him.

But, as I see it, the greatest value of graphics is in making a quick preliminary analysis of a problem to determine which variables to include and what kinds of mathematical functions to use. Here the econometrician is guided both by logical theory and by empiricism. Graphics can help him think through the problem. It can also help him choose functions that really describe the data. Then he may well choose to compute a mathematical function by least-squares or otherwise.

This handbook presents some typical examples of diagrams that are useful. The collection is, of course, far from complete. In fact, no collection will ever be complete, because new kinds of diagrams and

---

4/ See Ezekiel, Mordecai. Methods of Correlation Analysis. Ed. 2., 531 pp., illus. New York. 1941.

5/ See, for example, U. S. Dept. Agr. Tech. Bulletins 1068, p. 29; 1070, p. 10-13; 1080, p. 28, 52, 63; and 1105, p. 41, 57, 75, 89.

modifications of old diagrams are invented almost every day. Our purpose is only to encourage econometricians to experiment with graphics--not as their only tool of research, but as one that will help them do better work.

Like any other scientist, the econometrician tries to understand relationships between variables. He tries to explain variations in production, distribution, costs, prices, and profits. He cannot do this by logic alone. Pure theory is not enough. Nor does he find the answer by stating economic theory in mathematical terms which he can manipulate according to rules. Econometrics is the measurement of economic relationships. Theory must be tested and quantified.

To test and quantify economic theory, we must work with statistics. Often we can use statistics that are published regularly by public agencies or by private industry. Sometimes we will need special statistics obtained from surveys or from experiments. In any case, statistics are essential. But raw statistics alone are not worth much--just as pure theory alone is of little value. Theory and statistics must be combined if the research is to be worthwhile. Graphic analysis can be useful in this combination--especially in the preliminary stages.

Yields of Corn in Iowa, 1866-1943

The economist studies many kinds of variables. He is concerned with variations in prices, incomes, rates of consumption, production, and many other things. Ordinarily the economist wants to explain why certain things have varied in the past and what sort of changes can be expected in the future. Most of the diagrams in this report illustrate how to go about explaining why certain things have varied. However, we start with a diagram which shows simply how much variation there has been in the past without explaining why.

We know, of course, that much of the variation in yields of corn has been due to weather. We haven't learned as yet to control the weather. We may, however, expect that weather conditions in the future will vary about as they have in the past. With this in mind, what variations can we expect in yields of corn?

The diagram on page 5 6/ shows the frequency distribution of yields of corn in Iowa in the period 1866-1943. To make the data more nearly comparable over this long period, yields in recent years were adjusted downward to allow for estimated effects of the use of hybrid seed. This was a major factor in increasing yields during the late 1930's and the early 1940's.

The shaded area in the diagram (sometimes called a histogram) shows the actual number of years in which the yield was from 15.0 to 19.9 bushels, from 20.0 to 24.9 bushels, and so on. The irregularities in the steps would probably have differed slightly if data for fewer years or for other years had been shown. We can get a better idea of what we might expect generally by drawing a smooth curve such as the one shown on the diagram. Such curves can be computed mathematically from well-known formulas. In this case, the graphics indicate that the curve is skewed to the left. The analyst can easily see that a normal curve will not describe the variation. This is often the case with economic data. The statistician cannot safely assume that his frequency distributions are normal. He would do well to draw freehand graphic curves before trying any sort of mathematical fit. In many cases the graphic curve will satisfy all needs. When comparing several distributions and in certain other cases, mathematical curves and coefficients have merit.

As the more recent data used in this chart were adjusted downward to allow for the effect of the use of hybrid seed, the average level of yields suggested by the chart is not applicable currently. For example, yields per harvested acre in 1943-52, a period when nearly all of the corn grown in Iowa was from hybrid seed, averaged 50 bushels per acre compared with a yield of 40 bushels in 1920-29, a period having equally normal weather but when practically no hybrid seed was used. The chart suggests an average close to that in the 1920-29 period. The chart can be used, however, to indicate the probable variation in yield around some appropriate average.

6/ Unless otherwise specified, diagrams referred to in the text are those on the facing page, the data for which are given in the table beneath the chart.

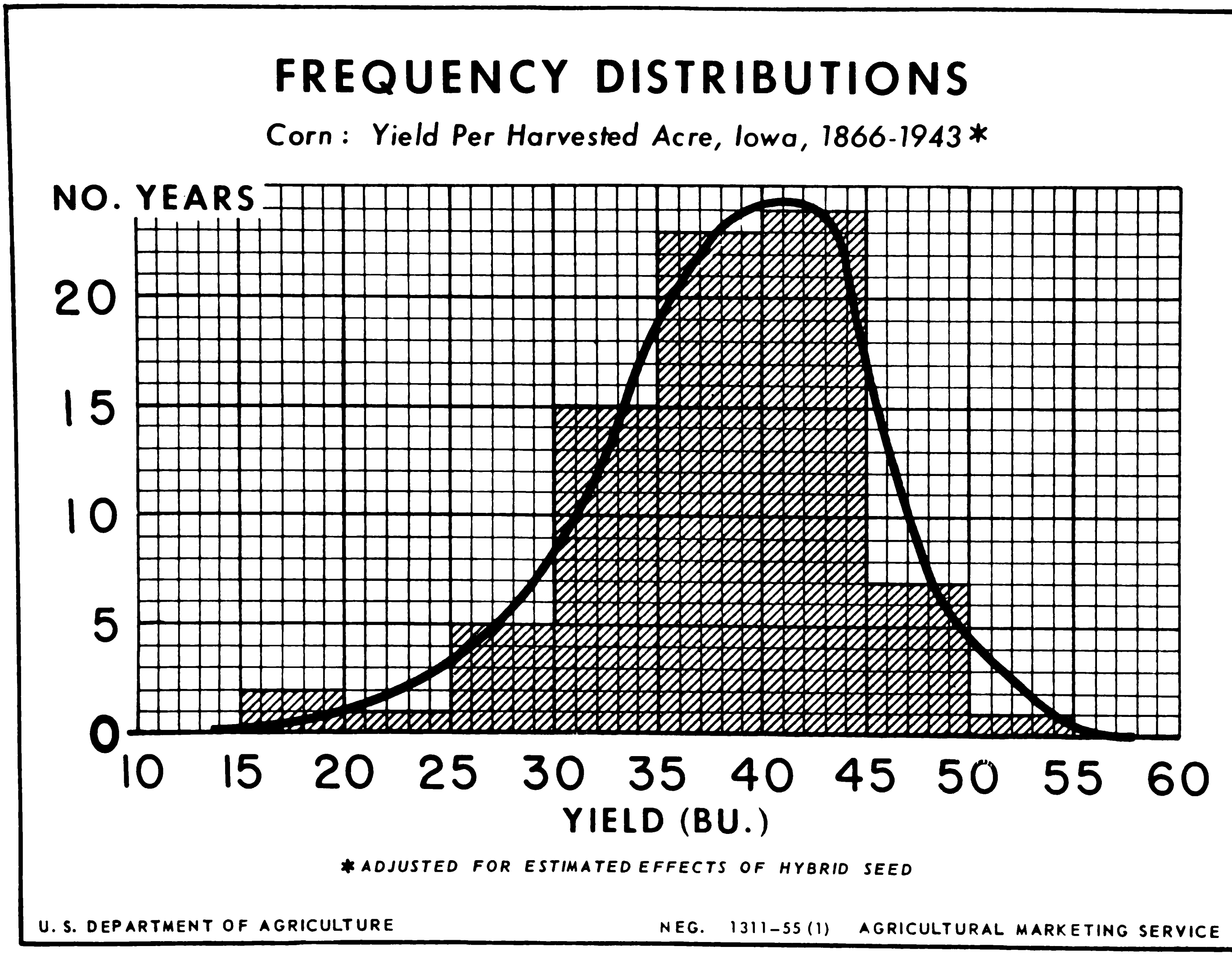


Figure 1

Corn: Frequency distribution of years classified by yield per harvested acre, Iowa, 1866-1943

Yield per harvested acre	Frequency of years
Bushels	Number
15.0-19.9	2
20.0-24.9	1
25.0-29.9	5
30.0-34.9	15
35.0-39.9	23
40.0-44.9	24
45.0-49.9	7
50.0-54.9	1

Foote, Richard J., and Bean, Louis H. Are Yearly Variations in Crop Yield Really Random? Agr. Econ. Research. 3:23-30, illus. 1951.

CUMULATIVE FREQUENCIES

Percentages of All Families and Unattached Individuals With Incomes Below Stated Levels

Economists and statisticians are concerned with many kinds of frequency distributions. The particular distribution shown on the diagram refers to percentages of families with various incomes. In this case we have shown a cumulative frequency curve. Thus, instead of showing the percentage of families with incomes from 0 to $1,000, from $1,001 to $2,000, etc., we have shown the percentage with incomes below $1,000, below $2,000, etc. This sort of cumulative frequency curve for incomes is sometimes known as a Lorenz curve.

The cumulative frequency curve, or ogive, has some advantages over the more usual noncumulative frequency curve. It can be used whatever the "class intervals" may be. For example, in this case, class intervals of $1,000 were used for the part of the curve from 0 to $5,000. For incomes above $5,000, the class interval was $2,500. With unequal class intervals, it is awkward to draw and use the ordinary type of frequency chart, and the cumulative chart is preferred.

In this case there was no problem of drawing a freehand curve to fit the cumulative frequencies. The plotted data all lie almost exactly along the freehand line we have drawn.

Several mathematical functions have been proposed and used to describe the distribution of incomes. Some of these, like the Pareto curve, are purely empirical. Others, like the Gibrat curve, are based upon logical considerations. It is obvious that no mathematical curve could fit the data much better than the freehand curve we have drawn. In fact, the freehand curve probably fits the data on the left hand side of the diagram better than would a mathematically fitted Pareto curve.

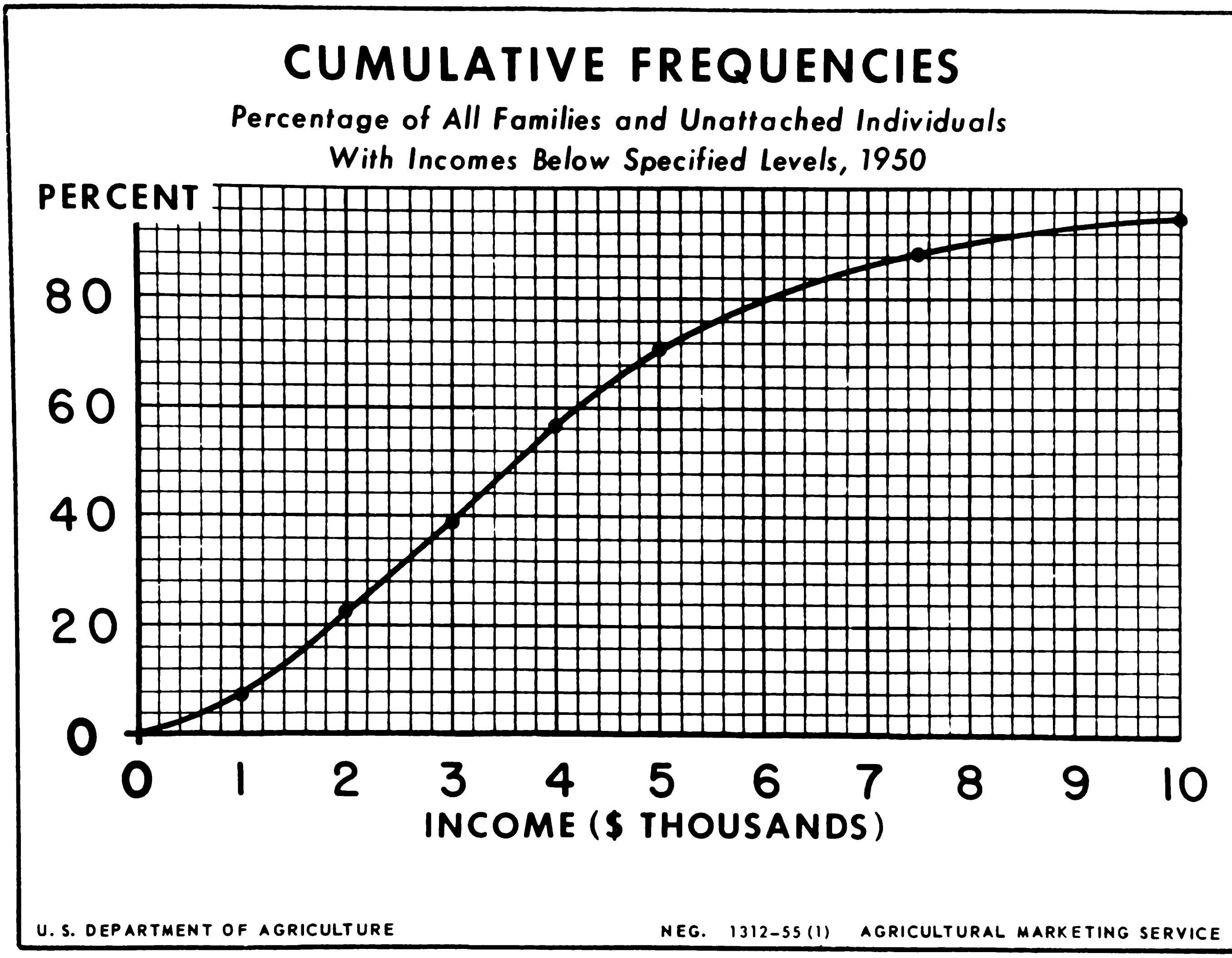


Figure 2

All families and unattached individuals: Percentage with incomes below specified levels, United States, 1950

Income level	Distribution
Dollars	Percent
Under:	
1,000	7.6
2,000	22.7
3,000	39.2
4,000	56.6
5,000	71.0
7,500	88.5
10,000	94.4

Goldsmith, Selma, et. al. Size Distribution of Income Since the Mid-thirties. Review Econ. and Statis. 36:4. 1954.

Farms Ranked by Total Value of Products Sold

We have just discussed the Lorenz curve. We now consider a somewhat different type of curve which is often useful.

The figures given below the diagram show that the lowest 20 percent of the farms got 0.4 percent of the income, the lowest 40 percent got 3.4 percent, and so on. These figures are plotted on the diagram with a smooth curve drawn through them that gives an estimate of the percent of the total value of farm products obtained by any given percent of the farms. For example, the lowest 95 percent of the farms received about 68 percent of the income. We can turn this around (reading the diagram upside down) and say that the top 5 percent of the farms got about 32 percent of the income.

On a diagram of this kind, if all farms had received the same income, the observation would have fallen along the straight dotted line. Thus, the degree of curvature is a measure of the inequality of income distribution.

Many mathematical functions have been used to study cumulative frequencies of incomes and other economic variables. The Pareto curve and the Gibrat curve are well known. Such mathematical functions are especially useful when comparing several curves. But simple graphics will help in any case to choose an appropriate function.

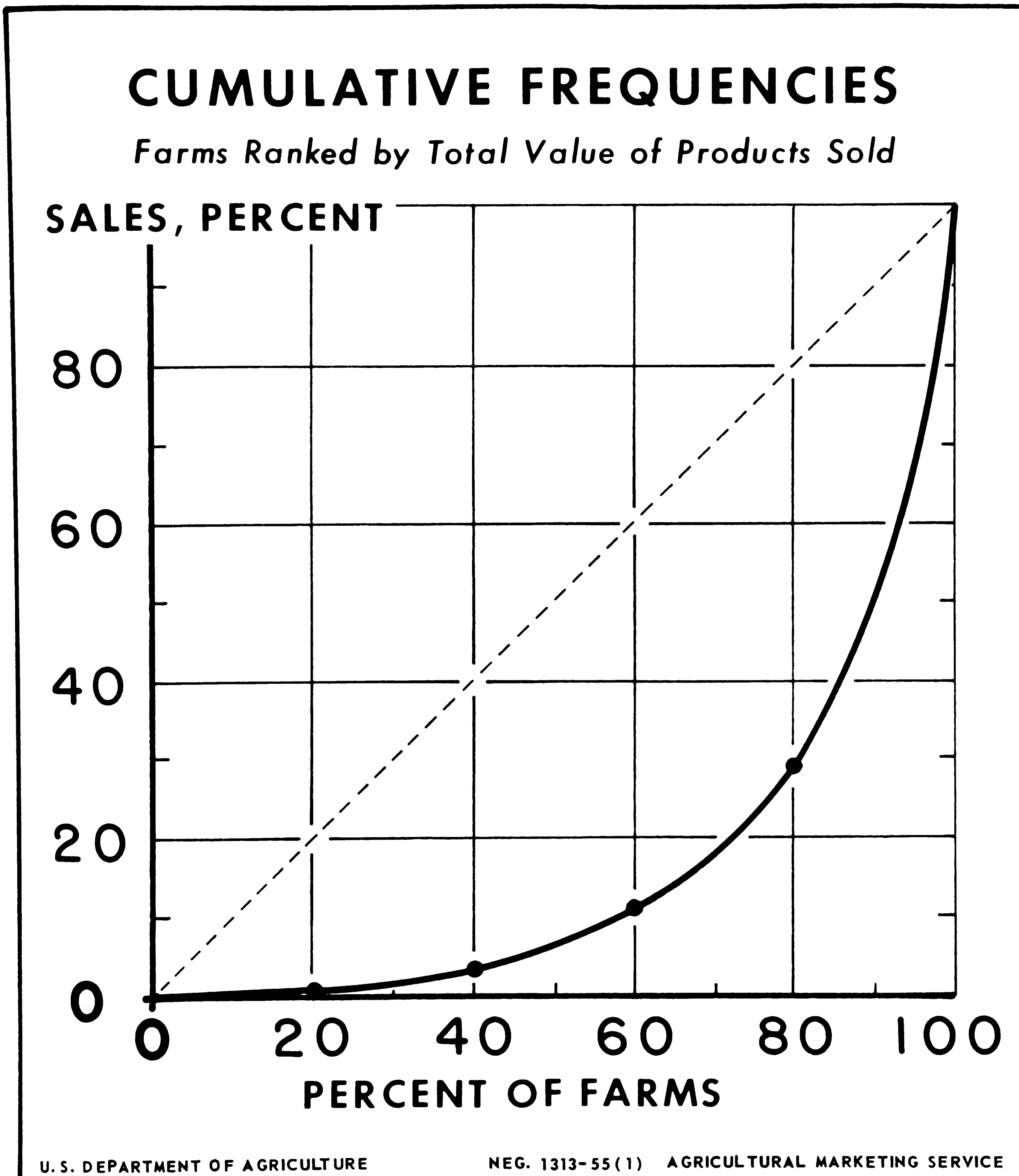


Figure 3

Relationship between percentage of farms and percentage of total value of products sold

Farms ranked by value of products	Percentage of total value of products sold
Percent	Percent
Lowest:	
20	0.4
40	3.4
60	11.2
80	29.2

Compiled from United States Census of Agriculture, 1950.

Population Trends in the United States by Decades, 1800-1950

The economist is often concerned with time trends. He wants to find out how some variable has been increasing or decreasing over a period of several years or decades. For example, he may be studying the growth of population in the United States or the rate of decline in the number of farm workers. In such cases he will want to disregard minor fluctuations due to errors in the data or to temporary disturbances. He will also generally want to disregard cycles or other shorter term movements in the data if they exist. He is concerned only with the gradual rate of change in a variable in relation to time.

The chart shows the Bureau of the Census estimates of the population in this country by decades since 1800. It is a simple matter to draw a freehand curve describing the trend. Ordinarily, at least, population does not change abruptly except by major wars, serious epidemics, or a sharp increase in immigration. If we plot the data for each year, or decade, we can usually draw a smooth curve running nearly through the points we have plotted. In this case, departures from the curve could well be due to errors in estimating the population. It should be noted that even official estimates may not warrant the naive faith in their accuracy that sometimes prevails. We, as economists, are probably as much responsible as any other group of users of published data for the insistence upon the publishing of a single number (point estimate) to represent, say, the population of the United States. We are reluctant to accept a lower and upper estimate (interval estimate) of the actual population even though we know that the Bureau of the Census official figure of 150,697,761 persons for 1950 (or any other year) may not be exact. All too often we do not take even the trouble to understand what the publisher has to say about the known, or estimated, amount of possible error in his estimates.

Instead of drawing a freehand curve, the statistician could, of course, fit some kind of mathematical function, such as a logistic curve. Our advice would be to draw a freehand curve first. In this case, it is doubtful if any mathematical function would give a better description of the trend than our freehand line. A mathematical curve might have some advantage when comparing trends in population in several different countries. If the same type of function was fitted in each case, results could be summarized in a few statistical measurements.

A practical application of trends is in forecasting. This always involves an extrapolation beyond the range of the data. Extrapolation of trends is dangerous whether it is done from a freehand curve or from a curve that has been fitted mathematically. For example, before the 1950 census data were available (so that we did not have the last observation on the diagram), many population experts drew an S-shaped curve indicating that the rate of growth had started to flatten. When this type of curve was extrapolated it suggested that the population would become stationary, or even decrease, by 1960 or 1970. Such an extrapolation now looks doubtful in view of the census figure for 1950.

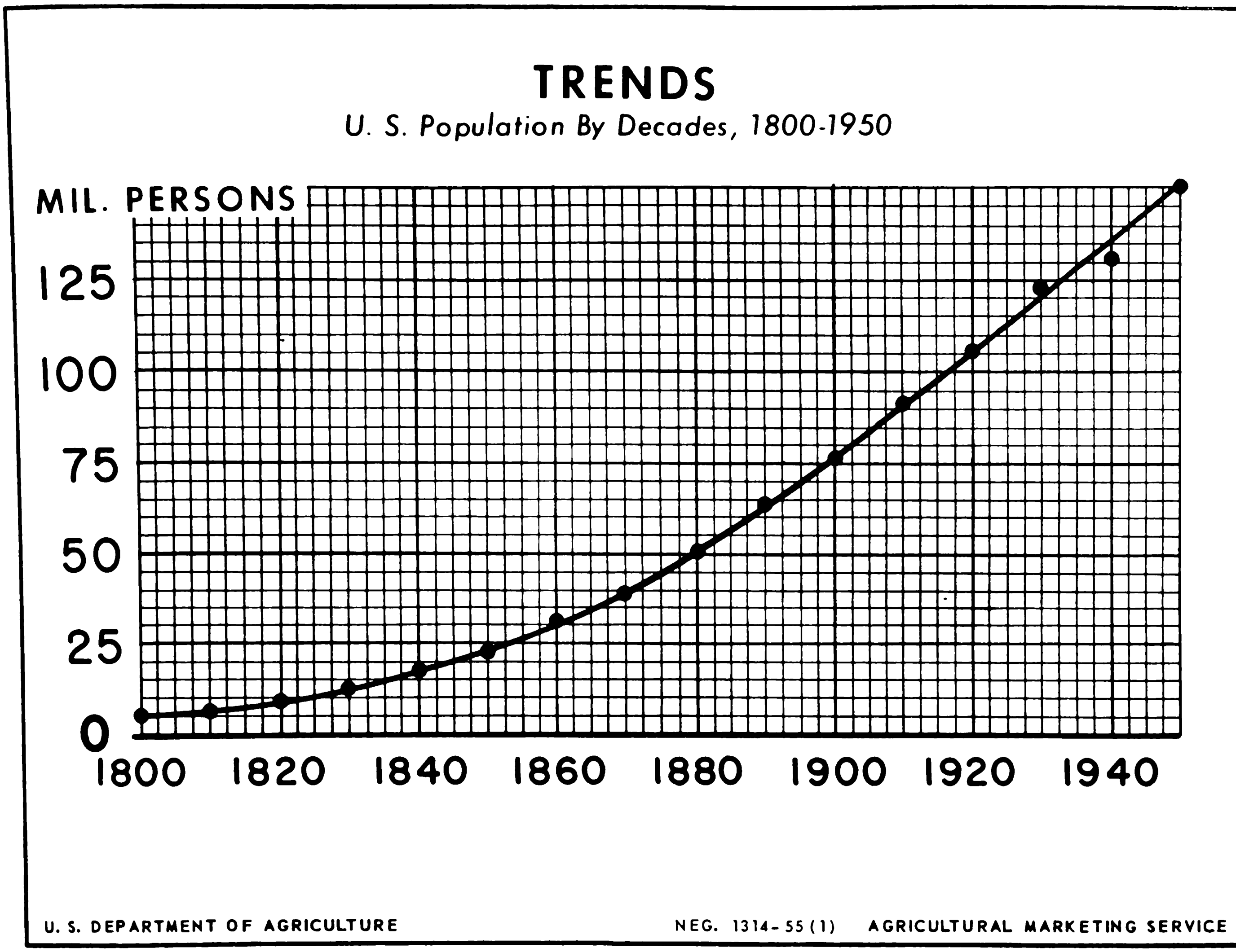


Figure 4

Population: United States, by decades, 1800-1950

Year	Population	Year	Population
	Millions		Millions
1800	5.3	1880	50.2
1810	7.2	1890	62.9
1820	9.6	1900	76.0
1830	12.9	1910	92.0
1840	17.1	1920	105.7
1850	23.2	1930	122.8
1860	31.4	1940	131.7
1870	38.6	1950	150.7

Bureau of the Census.

Volume of Agricultural Marketings

In most cases, of course, successive observations by years or decades when plotted on a chart do not lie exactly upon a smooth line. However, it is usually possible to draw a freehand line or curve describing the general trend.

This diagram shows an index of the volume of agricultural marketings each year from 1910 through 1953. Obviously, the growth has not been so steady as the growth of population. For example, marketings did not increase from 1928 through 1937. This was due in part to a business depression and low prices, and in part to two serious droughts. During World War II, there was a remarkable expansion in agricultural marketings to meet the needs of domestic and foreign markets. After 1945, there was a slow increase. We need a trend line that will describe these characteristics. The freehand line shown on the diagram perhaps describes them fairly well.

Again, the statistician may want to use some mathematical function to describe this trend. Before doing so he would be well advised to draw first a freehand trend in order to indicate what sort of function should be used. In this case, for example, a straight line would not adequately describe the observed trend. A third-degree parabola might give a fair fit, but would be an extraordinarily bad curve to extrapolate into the future. In general, the economist would do well to avoid parabolas. Logically, they seldom make any economic sense.

For most purposes, the freehand trend is as good as any mathematical trend we might compute. An exception might be the problem of comparing trends in several variables. Suppose we wanted to compare the trend in agricultural marketing with trends in agricultural output, the amount of fertilizer used, population, and so on. We might find some type of mathematical function that fits all the trends reasonably well. Then each curve could be summarized by a few constants. It would be easy to compare one with another. But this sort of procedure often would cover up some interesting and important features of some of the trends. It is always a good idea to plot the data and to draw freehand curves first.

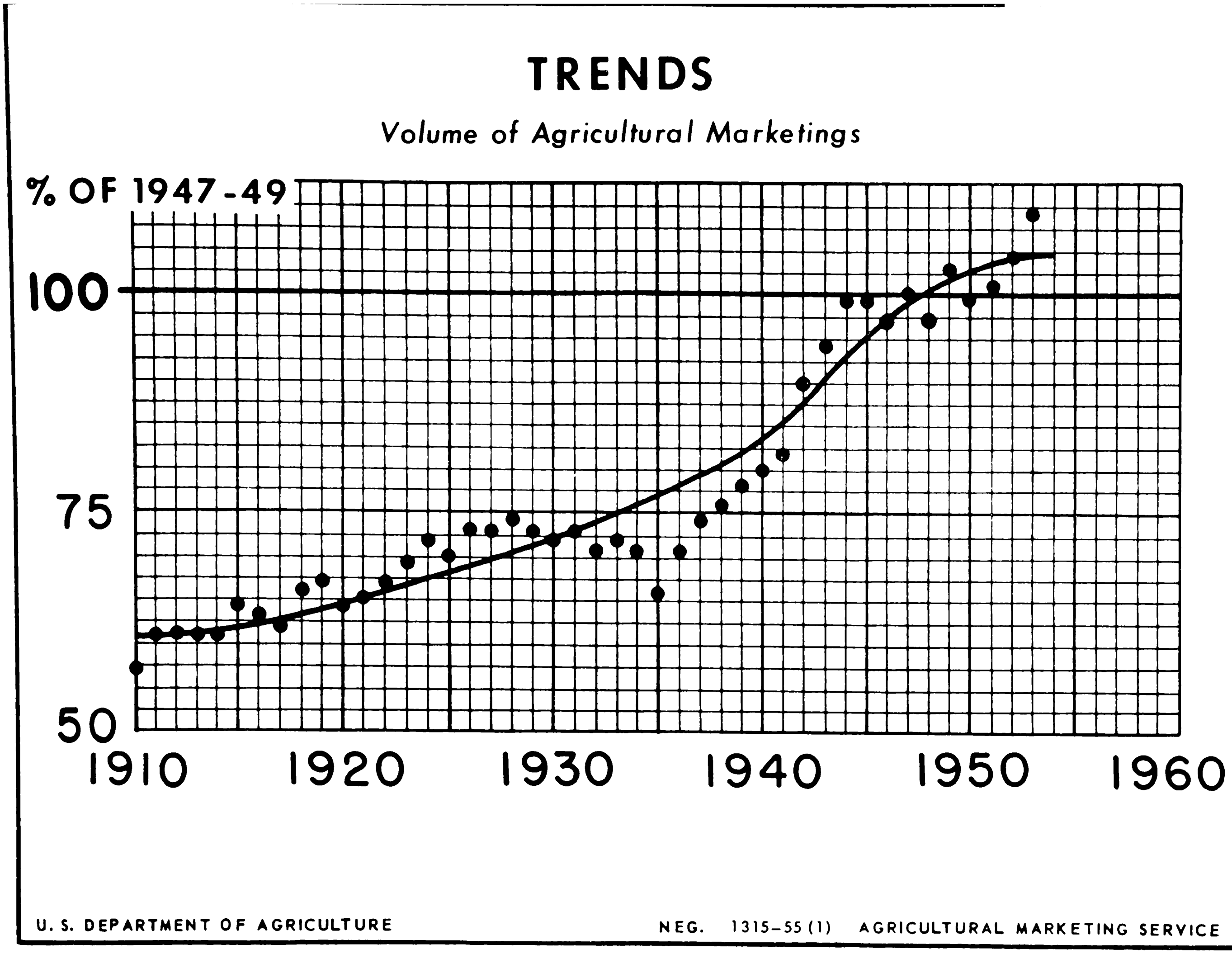


Figure 5

Farm marketings and home consumption: Index numbers of volume, 1910-53

[1947-49=100]

Year	Volume	Year	Volume	Year	Volume
1910	57	1925	70	1940	80
1911	61	1926	73	1941	82
1912	61	1927	73	1942	90
1913	61	1928	74	1943	94
1914	61	1929	73	1944	99
1915	64	1930	72	1945	99
1916	63	1931	73	1946	97
1917	62	1932	71	1947	100
1918	66	1933	72	1948	97
1919	67	1934	71	1949	103
1920	64	1935	66	1950	99
1921	65	1936	71	1951	101
1922	67	1937	74	1952	104
1923	69	1938	76	1953	109
1924	72	1939	78		

Agricultural Marketing Service.

CYCLES

Cattle on Farms, January 1

Some economic data exhibit successive cycles of approximately the same length covering periods of up to several years. This is especially true in the case of some agricultural data such as those on cattle numbers. When cattle prices are high, farmers are likely to start breeding for larger herds. It takes several years to increase the herds substantially, and the increase ordinarily continues for some time after prices become unprofitable. Then the reverse happens and herds are gradually decreased.

When the annual data on cattle numbers are plotted, as shown in the diagram, it is easy to see that there have been fairly regular ups and downs. We have drawn a smooth curve through the data to describe these ups and downs.

The statistician could, of course, fit some kind of mathematical function to data of this kind. He would choose a curve that would allow for an upward trend, and for the cyclical swings around the trend. But it would take a very complicated mathematical curve to fit the data as well as the freehand curve.

The main practical interest in cycles stems from the need for forecasts. The farmer naturally wants to know where we are in the current cycle--are cattle numbers approaching a peak and when are they likely to turn down? An analysis of past history will help him answer such questions. But he will do well to give special attention to current developments. For example, he will need to watch current trends in cattle slaughter.

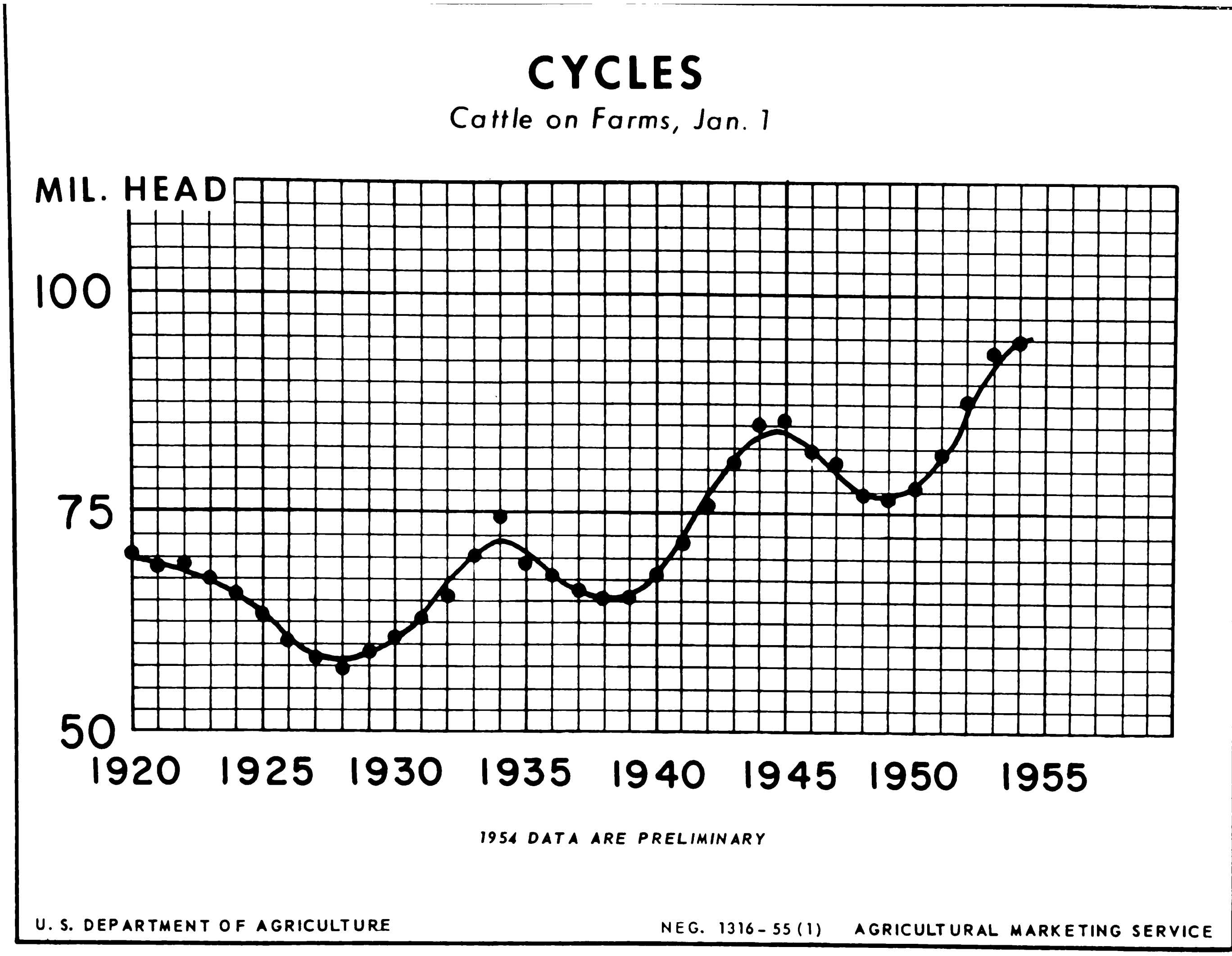


Figure 6

All cattle and calves: Number on farms January 1, 1920-54

Year	Number	Year	Number	Year	Number
	Millions		Millions		Millions
1920	70.4	1932	65.8	1944	85.3
1921	68.7	1933	70.3	1945	85.6
1922	68.8	1934	74.4	1946	82.2
1923	67.5	1935	68.8	1947	80.6
1924	66.0	1936	67.8	1948	77.2
1925	63.4	1937	66.1	1949	76.8
1926	60.6	1938	65.2	1950	78.0
1927	58.2	1939	66.0	1951	82.0
1928	57.3	1940	68.3	1952	87.8
1929	58.9	1941	71.8	1953	93.6
1930	61.0	1942	76.0	1954	1/94.7
1931	63.0	1943	81.2		

1/ Preliminary.

Agricultural Marketing Service.

Cattle on Farms by Cycles

One of the best ways to forecast the probable behavior of a current cycle from that of previous cycles is to break the total series into individual cycles. In 1954, we were about halfway through the most recent cycle in numbers of cattle on farms. In the diagram shown here, data for these individual cycles are plotted on the same scale, beginning with the year of the low point in inventories in each instance.

The several cycles of numbers of cattle are remarkably similar. One handicap in this visual scheme is that each cycle is of a different length. Similarity between cycles would appear even closer if the cycles were telescoped into a uniform length.

A good statistician knows that history seldom repeats itself exactly. Cycles vary in length and in amplitude. A knowledge of past trends, and of past cycles, gives some perspective to the present. Often it suggests the general direction of changes in the immediate future. But the wide-awake economist will be looking for factors that may make the current cycle different from the others.

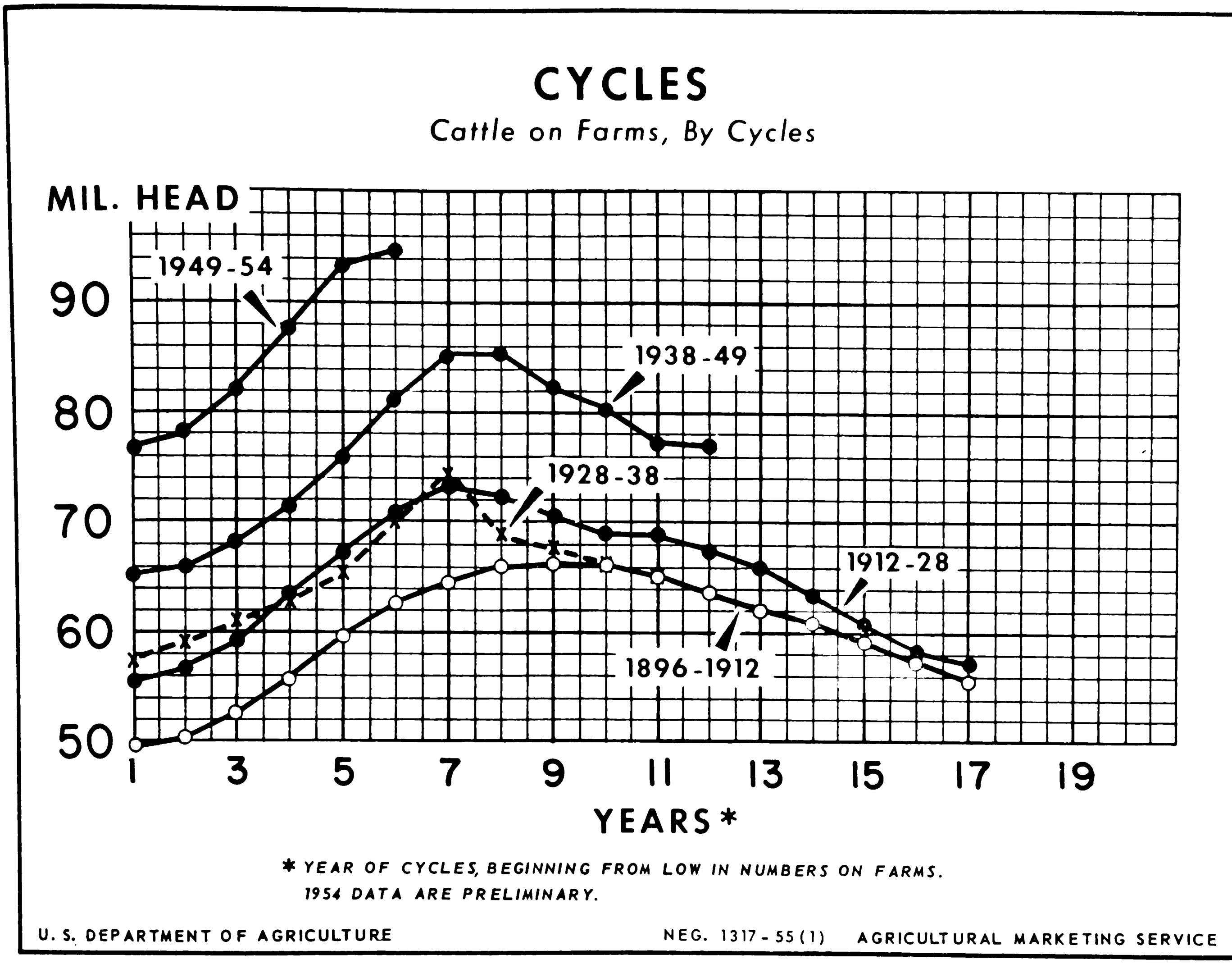


Figure 7

All cattle and calves: Number on farms January 1, 1896-1919 1/

Year	Number	Year	Number
	Millions		Millions
1896	49.2	1908	62.0
1897	50.4	1909	60.8
1898	52.9	1910	59.0
1899	55.9	1911	57.2
1900	59.7	1912	55.7
1901	62.6	1913	56.6
1902	64.4	1914	59.5
1903	66.0	1915	63.8
1904	66.4	1916	67.4
1905	66.1	1917	71.0
1906	65.0	1918	73.0
1907	63.8	1919	72.1

1/ See tabulation on page 15 for data for later years.

Agricultural Marketing Service.

Hog Slaughter and the Hog-Corn Price Ratio

The agricultural economist usually is not content with simply observing periodic movements in prices or in production. He wants to know what causes the swings. And he especially wants to know how the current cycle is developing--whether, for example, it will be shorter or longer than average.

Of course, if all cycles were completely regular, all the statistician would have to do is to find some kind of curve, making allowance for trend and for cyclic ups and downs. A projection of this curve would be a forecast of what is likely to happen in the next few months or years. The business-cycle analysts have learned from painful experience that such mechanical forecasts are unreliable. Many of the cycles in agriculture are more regular than those in business. Still there is a good deal of variation in agricultural cycles. To understand what is going on and what is likely to happen in the immediate future, the agricultural economist analyzes the forces that have shaped the cycles in the past and that appear to be influencing the current cycle.

The accompanying chart illustrates a simple analysis of the hog cycle. The lower part of the chart shows the number of hogs slaughtered each year for 1920-54. The upper part of the chart shows the ratio between hog prices and corn prices (that is, the number of bushels of corn required to buy 100 pounds of live hogs, based upon average prices received by farmers for hogs and corn). A high hog-corn ratio indicates the situation in which hog production is profitable; a low ratio indicates the opposite.

By comparing the two parts of the chart it is easy to see that changes in hog slaughter lag a year or two behind the hog-corn ratio. The dotted lines connect some of the peaks and troughs of the two curves.

By keeping a chart of this kind up to date, an economist can get a fairly accurate idea of probable developments in the next six-months-to-a-year period.

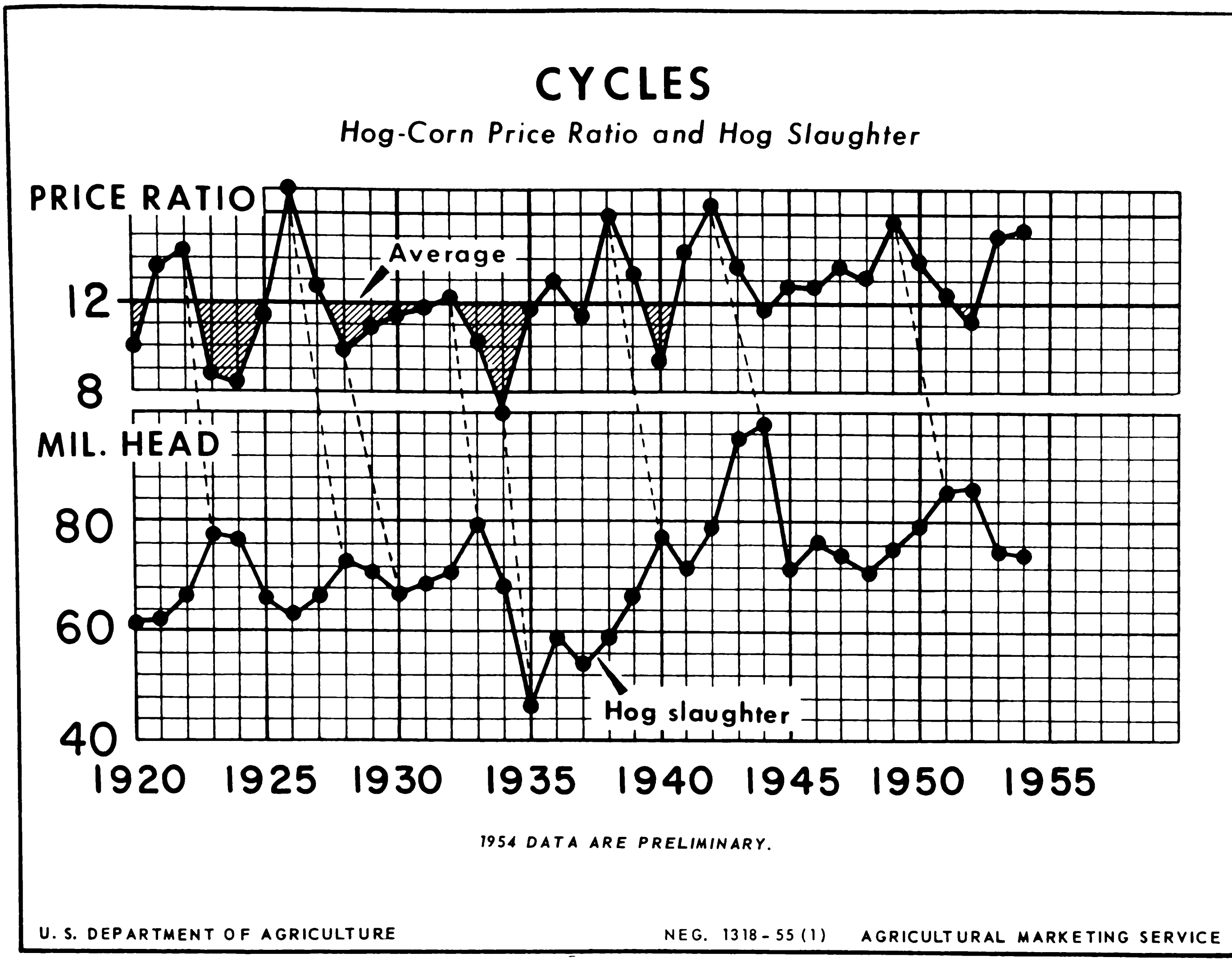


Figure 8

Hogs: Number slaughtered and hog-corn price ratio, 1920-54

Year	Slaughter of hogs	Hog-corn price ratio 1/	Year	Slaughter of hogs	Hog-corn price ratio 1/	Year	Slaughter of hogs	Hog-corn price ratio 1/
	Millions			Millions			Millions	
1920	61.5	9.8	1932	71.4	12.3	1944	98.1	11.6
1921	61.8	13.6	1933 2/	79.7	10.4	1945	71.9	12.8
1922	66.2	14.4	1934	68.8	7.0	1946	76.1	12.6
1923	77.5	8.7	1935	46.0	11.6	1947	74.0	13.6
1924	76.8	8.2	1936	58.7	13.0	1948	70.9	13.0
1925	65.5	11.4	1937	53.7	11.1	1949	75.0	15.7
1926	62.6	17.0	1938	58.9	16.0	1950	79.3	13.7
1927	66.2	12.7	1939	66.6	13.3	1951	85.6	12.4
1928	72.9	9.9	1940	77.6	9.2	1952	86.7	11.0
1929	71.0	10.9	1941	71.4	14.2	1953	74.8	15.0
1930	67.3	11.4	1942	78.5	16.5	1954 3/	74.0	15.4
1931	69.2	11.7	1943	95.2	13.6			

1/ Number of bushels of corn required to buy 100 pounds of live hogs at local markets, based on average prices received by farmers for hogs and corn. Annual average is straight average of monthly ratios. 2/ Includes those slaughtered for Government account. 3/ Preliminary.

Agricultural Marketing Service.

SEASONAL VARIATION

Egg Prices, 1925-29 and 1945-49

Many economic time series follow rather regular seasonal patterns--moving in 12-month cycles. This is especially true of many agricultural time series. Production of some agricultural commodities naturally follows a regular seasonal pattern, which results from the seasonal pattern of the weather and certain production practices related to it. This tends to bring about an annual cycle in marketing and prices--especially in the case of perishables.

The mathematical statistician is sometimes tempted to make a mechanical analysis of seasonal variation. He may fit to the data a combination of a linear trend and a sine curve--and he may feel satisfied with a high correlation coefficient. But usually a simple graphic analysis will show up some important facts that might otherwise escape the researcher.

The first part of any analysis of seasonal pattern must be rather mechanical. We must fit some sort of trend and study the monthly deviations from it. The diagram is based upon an average of deviations of monthly egg prices from a 12-month moving average--an appropriate sort of trend for our purposes. We have shown the average deviation for each month of the year in two different 5-year periods.

The striking fact brought out by this diagram is that the seasonal pattern in egg prices is changing. Low prices still occur in the spring, high prices in the fall. But the seasonal swing is much less than it was 20 years ago. And the peak price comes earlier in the fall. These changes chiefly are a reflection of new and improved practices on the farm. The trend toward a less pronounced seasonal pattern and toward an earlier fall peak in egg prices is continuing. These facts are obviously important to farmers and to storers of eggs. Without a graphic analysis, such important facts could be easily overlooked. It is not impossible, of course, to fit a mathematical curve which allows for a damping of the seasonal swing. But the point is that a simple graphic analysis shows what sort of curve is needed.

A graphic method of measuring shifts in the seasonal pattern over time is illustrated in figure 11.

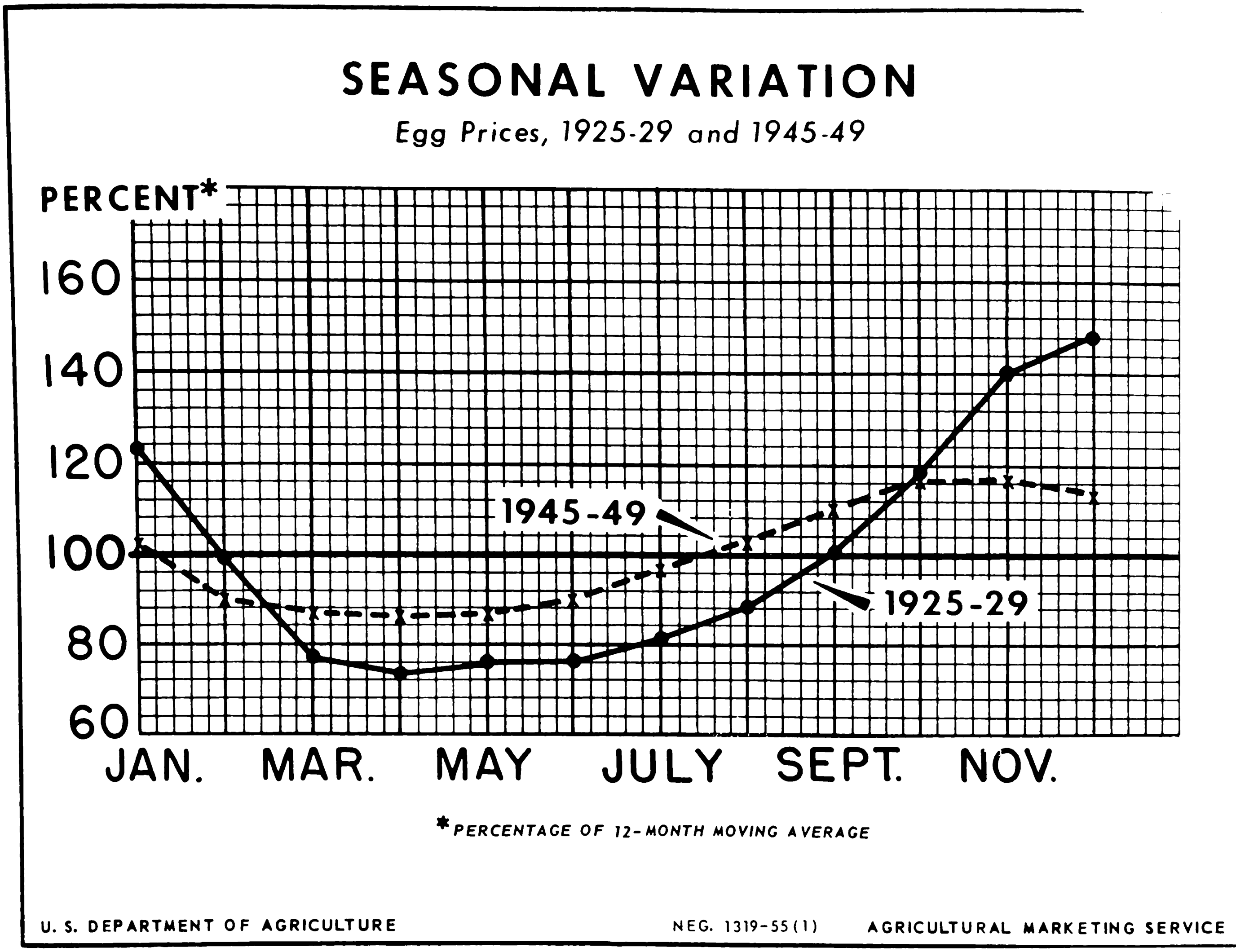


Figure 9

Eggs: Index numbers of seasonal variation of prices received by farmers, 1925-29 and 1945-49 1/

Month	1925-29	1945-49
January	123	102
February	99	90
March	77	87
April	73	86
May	76	87
June	76	90
July	81	97
August	88	103
September	101	110
October	118	117
November	140	117
December	148	114

1/ Average of percentages of 12-month moving average.

Agricultural Marketing Service.

Monthly Production of Pork and Prices Received by Farmers for Hogs

When discussing cycles, we observed that the economist usually wants to analyze the forces that have been responsible for upward and downward swings. The same is true with seasonal variation. Some fairly regular seasonal movements are well-known, such as the low price of eggs in the spring and the higher price during the late fall and early winter. However, no two years are just alike; sometimes the seasonal swing is big, sometimes it is little.

The accompanying chart shows the average seasonal variation in the production of pork and in prices received by farmers for hogs. The seasonal low point in prices comes in November and December when production is high. As pork production falls off in winter, prices go up somewhat. The swing in price is much less marked than the swing in production. This is due essentially to storage, which evens out the supply of pork to the consumer. In an average year, the price must rise enough in summer to cover storage costs.

This chart, of course, illustrates only the average seasonal variation for the postwar years. To be most useful as a guide to current marketing operations, we would need to break down these averages to show how the seasonal variation in hog prices is affected by different kinds of seasonal patterns in production. In that way we could forecast more accurately the seasonal price changes for a current year.

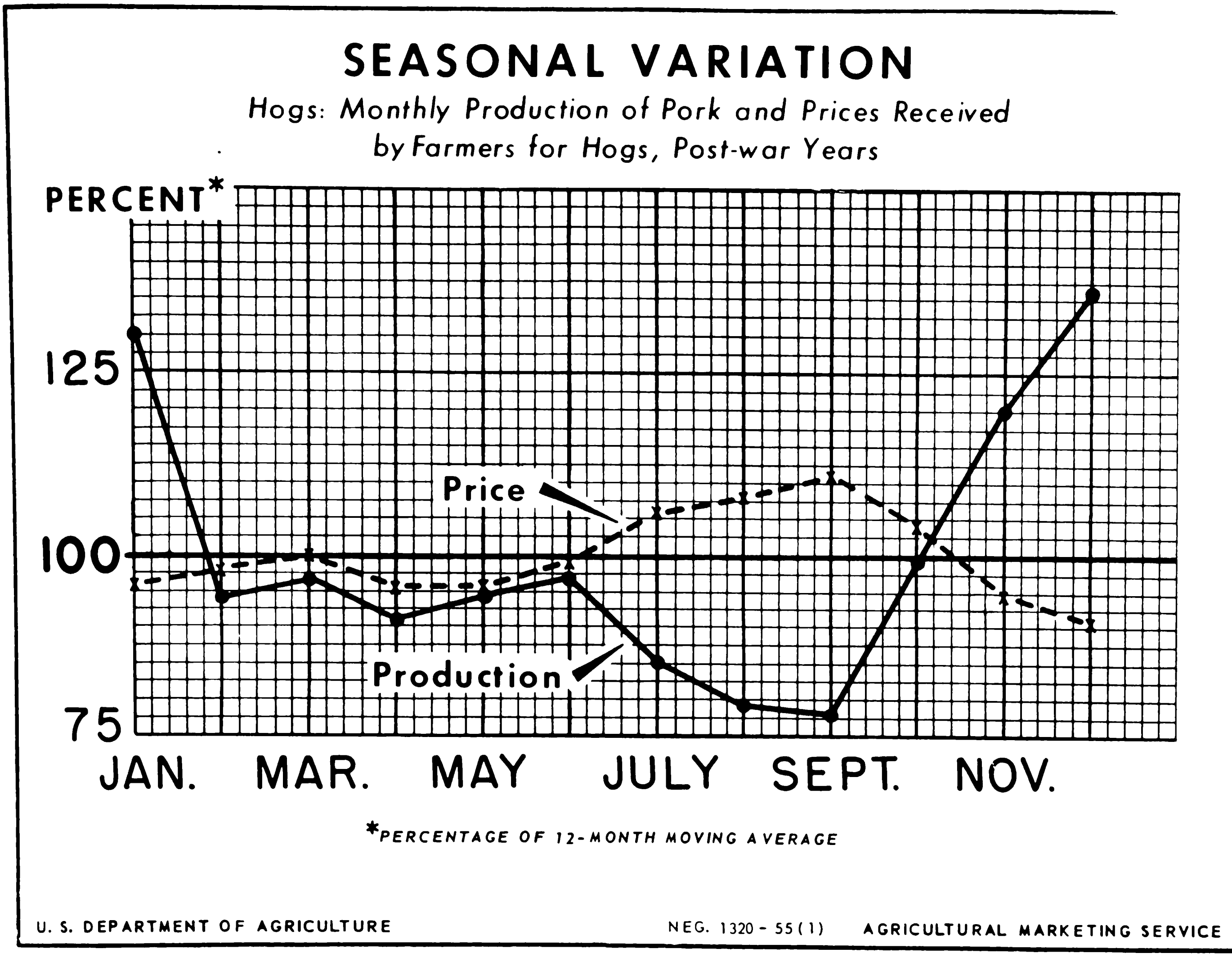


Figure 10

Hogs: Index numbers of seasonal variation of production of pork and price received by farmers for hogs, postwar years

Month	Production 1/	Price
January	130	96
February	94	98
March	97	100
April	91	96
May	94	96
June	97	99
July	85	106
August	79	108
September	78	111
October	99	104
November	120	95
December	136	91

1/ Excluding lard.

Breimyer, Harold F., and Johnson, Lucille W. Seasonality in Marketings and Prices of Meat Animals. U. S. Agr. Mkt. Service. Livestock and Meat Situation, Nov.-Dec. 1952, pp. 12-17, illus. (Processed).

Trends in the Seasonal Pattern for Eggs

A convenient way to determine graphically whether there has been a significant shift over time in the seasonal pattern is to plot the ratios to trend for all years against time in a series of charts, using a separate chart for each month. In most cases, visual inspection indicates whether the degree of seasonal variation is significant and whether it has changed during a period of time. If most of the values for any given month are consistently above or below a ratio of 1 by a fairly uniform amount for all years, it can be assumed that a seasonal pattern prevails and that it has not changed significantly during the period. If the charts do not indicate clearly whether the seasonal pattern is significant and if no change over time in the pattern is indicated, a mathematical test based on analysis of variance described in Foote and Fox 7/ can be applied.

Such a chart for eggs for December is shown in this diagram. In this it is clear that the nature of the seasonal pattern has shifted over time, thus verifying the findings from the chart on page 21. In this case, a linear trend might fit the ratios fairly well. If a linear trend appeared to be applicable on the other 11 charts, the 12 trends could be fitted simultaneously by a mathematical method described by Foote and Fox. The advantage of this over a graphic fit is that the trends are fitted subject to the condition that the constant values add to 1,200 and the regression coefficients add to zero. Hence, the computed index numbers of seasonal variation for the 12 months for each year add to 1,200. If a curvilinear or irregular trend on at least some of the charts is indicated, all of the trends can be fitted graphically in such a way that the trend values for the 12 months in each year add to approximately 1,200.

7/ Foote, R. J., and Fox, Karl A. Seasonal Variation: Methods of Measurement and Tests of Significance. U.S. Dept. Agr. Agr. Handb. 48, 16 pp. 1952.

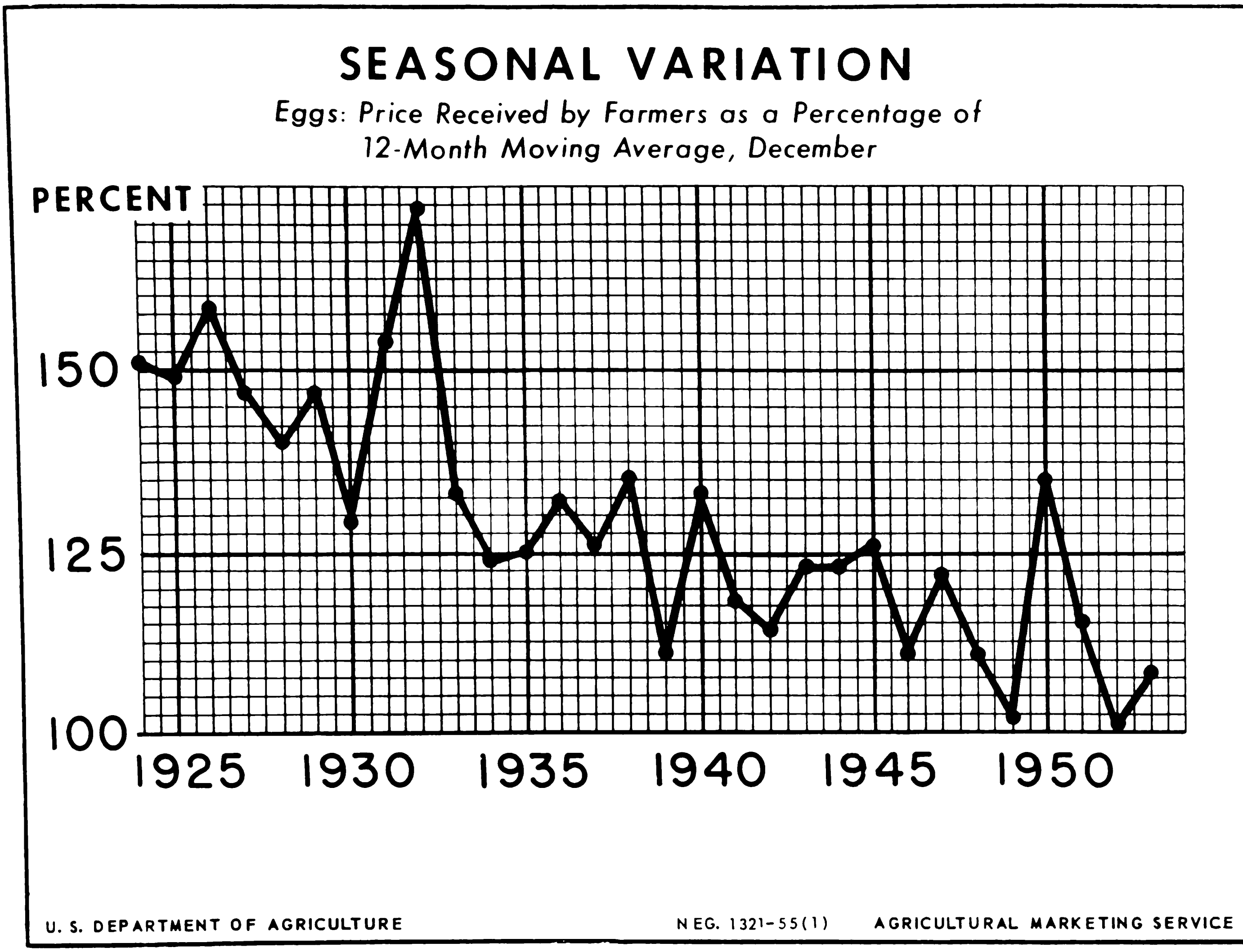


Figure 11

Eggs: Price received by farmers as a percentage of
12-month moving average, December, 1924-53

Year	Percentage	Year	Percentage
	Percent		Percent
1924	151	1939	111
1925	149	1940	133
1926	158	1941	118
1927	147	1942	114
1928	140	1943	123
1929	147	1944	123
1930	129	1945	126
1931	154	1946	111
1932	172	1947	122
1933	133	1948	111
1934	124	1949	102
1935	125	1950	135
1936	132	1951	115
1937	126	1952	101
1938	135	1953	108

Agricultural Marketing Service.

Relation of Corn Yields to Nitrogen

The so-called "dot chart" is one of the handiest tools of economic analysis. Any competent economic analyst draws dot charts and studies them before putting numbers in a calculating machine.

This is a fairly typical example of a dot chart. The construction of the chart is easy. For example, the data obtained by certain experiments in North Carolina indicate that with 20 pounds of nitrogen per acre, an average yield of 41.5 bushels of corn is obtained. To plot this observation we simply measure 20 units along the horizontal axis and then 41.5 units upward and mark that point with a dot. Similarly, for the other points on the diagram.

Through these points we draw a smooth curve to indicate the response of corn yields to varying amounts of nitrogen. In this case this is easy to do because all of the observations (that is, all the dots) lie close to a smooth curve. This is because the data are averages obtained from controlled experiments. If we had used individual data obtained from the survey covering a number of farms in different parts of the State, with varying soil and moisture conditions, the dots probably would not have clustered so closely around the smooth curve.

Economists are, of course, particularly concerned with input-output relationships. In agriculture a great deal of research has been done on such matters as the response of crop yields to fertilizer application and on the response of animals to varying amounts and kinds of feeds. Mathematicians have suggested certain mathematical formulas to measure such responses. For example, specific formulas have been suggested by Mitscherlich, Spillman, Cobb, and Douglas. These mathematical formulas are all based upon logical considerations and are of considerable interest to economists. Still, none of them may show accurately the relationship which can be seen easily in the dot chart. The economist should not forget logical considerations in drawing freehand lines or curves. The curve we have drawn in this case probably represents the relationship at least as satisfactorily as any of the mathematical functions. In addition, there may be some question about the logic of any of the proposed mathematical functions when they are extrapolated far beyond the range of the data. For example, none of the three mathematical functions mentioned would allow for the possibility that excessive amounts of nitrogen might actually reduce corn yields.

In passing we might note that the curve indicates both decreasing average returns and decreasing marginal returns throughout the range of observations. The graphic derivation of marginal curves is described on page 56 .

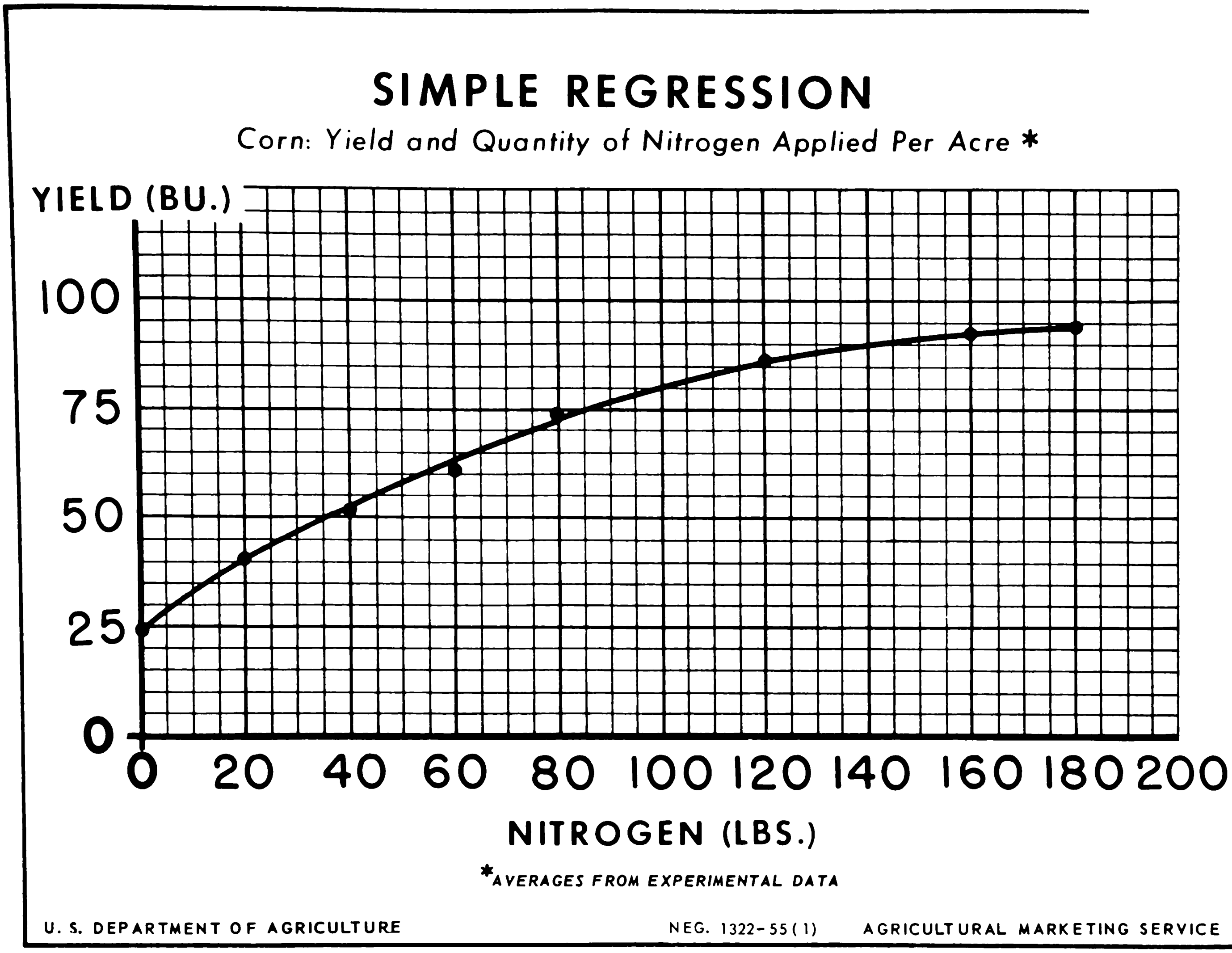


Figure 12

Corn: Yield per acre by specified quantities of nitrogen applied

Nitrogen applied	Yield of corn
Pounds	Bushels
0	24.5
20	41.5
40	52.1
60	61.4
80	73.5
120	86.0
160	92.8
180	94.0

Johnson, Paul R., Alternative Functions for Analyzing a Fertilizer-Yield Relationship. Jour. Farm Econ. 35:519-529. 1953.

Food Expenditures Related to Incomes Based on Averages from Survey Data

The first chart on simple regression used averages obtained from experimental data. In cases of this kind, it may be possible to control most of the factors affecting the dependent variable. Because of this control, the observations may lie close to a smooth curve.

The economist can seldom make experiments. He usually has to rely on data obtained from surveys or from published time series.

This chart shows the relationship of food expenditures to per capita incomes as indicated by two surveys. Here again the observations cluster closely around the two smooth curves we have drawn. This does not necessarily indicate a high correlation between the food expenditures of individual families and the incomes available to those families. Actually there is a great variation in food expenditure within a group of families getting the same income. This variation is covered up in the averages. This is all right for our purpose, assuming that we want to estimate the average food expenditure for families with any given income. In this case, food expenditure is the dependent variable (that is, the variable we are trying to estimate).

When using survey data of the kind indicated here the economist relies heavily upon large numbers of observations. The two curves shown on the diagram appear to be reasonably accurate. This is indicated by the fact that the observations lie very close to the smooth curves. The analysis shows that each income group spent more for food in 1941 than in 1935.

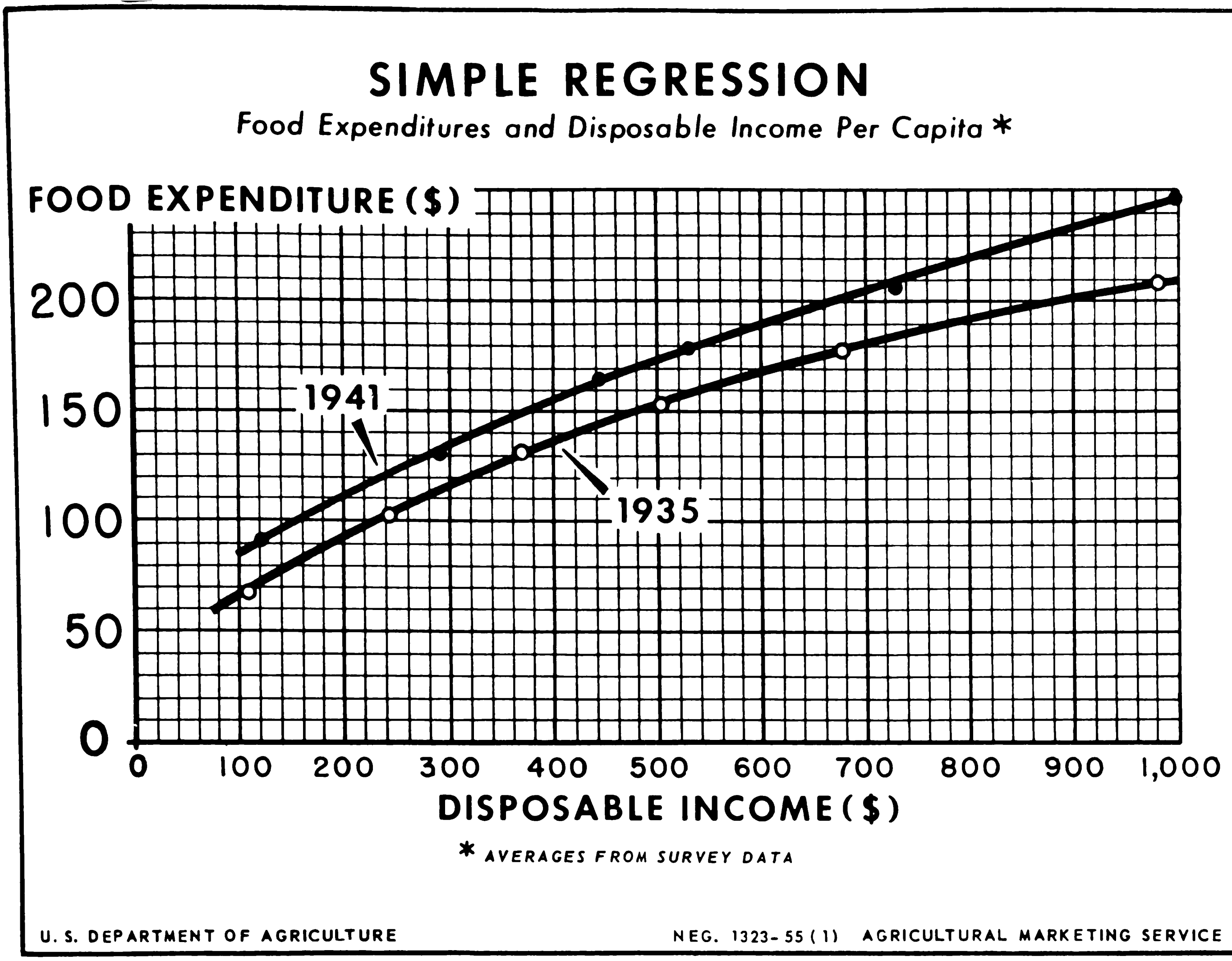


Figure 13

Expenditures for food, per capita, by specified income groups, 1935 and 1941

Disposable income			Expenditures for food	
Per consumer unit	Per capita			
	1935 1/	1941	1935 1/	1941
Dollars	Dollars	Dollars	Dollars	Dollars
Under 500	113	122	69	91
500 - 999	242	293	104	130
1,000 - 1,499	370	446	132	167
1,500 - 1,999	502	529	154	179
2,000 - 2,999	679	734	179	206
3,000 - 4,999	982	1,008	209	247
5,000 and over	3,270	2,027	344	354

1/ Most of the data used in income and expenditure studies of the National Resources Committee relate to year beginning July 1935. Some data, however, cover calendar year 1935. Data shown here were derived from the studies.

Burk, Marguerite C. A Study of Recent Relationships Between Income and Food Expenditures. Agr. Econ. Research. 3:87-97, illus. 1951.

Onion Prices Related to Production

Section A of the chart in this example is a dot chart showing the relation between onion production and prices in the years 1939-52. You will note that the observations are scattered all around the diagram and that some of the highest prices occurred in the years of medium to large production. Also, some of the lowest prices occurred in years of low production.

This does not indicate a positively sloping demand curve. It indicates only that both prices and production increased during the period studied. To get a rough idea of the relation between production and prices, we have drawn a line from each observation to each succeeding observation. This is generally a good practice in dealing with time series. It quickly shows up any trend in the data and gives a rough idea at least of the slope of the curve.

In this particular case, section A suggests that we consider the relation of year-to-year changes in prices and in production. This relation is shown in section B. It appears that changes in production give a fairly good indication of expected changes in prices. The explanation is far from perfect. For example, if we had used the curve in section B to estimate expected changes in prices we would have been almost 40 cents too high in 1945 and about 45 cents too low in 1952.

A more accurate way of studying the relation between onion prices and production is discussed on page 40.

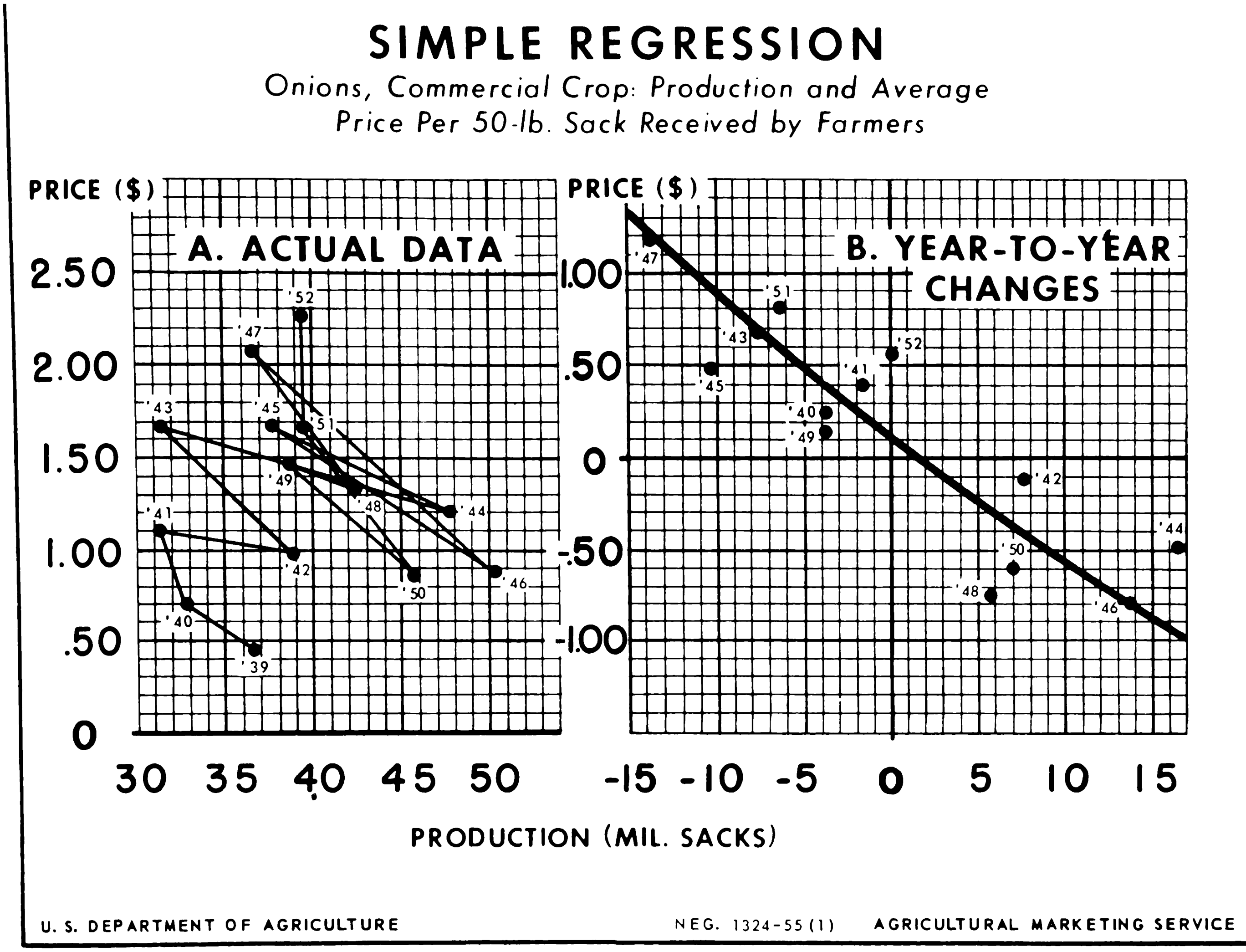


Figure 14

Onions, commercial crop: Production and average price per 50-pound sack received by farmers, 1939-52

Year	Production	Price	Change from preceding year in- Production	Change from preceding year in- Price
	Million sacks	Dollars	Million sacks	Dollars
1939	36.6	0.45	---	---
1940	32.9	.70	- 3.7	0.25
1941	31.2	1.10	- 1.7	.40
1942	38.9	.99	7.7	-.11
1943	31.3	1.68	- 7.6	.69
1944	47.9	1.20	16.6	-.48
1945	37.7	1.69	-10.2	.49
1946	50.4	.89	12.7	-.80
1947	36.7	2.08	-13.7	1.19
1948	42.5	1.32	5.8	-.76
1949	38.8	1.47	- 3.7	.15
1950	45.8	.87	7.0	-.60
1951	39.4	1.67	- 6.4	.80
1952	39.4	2.25	0	.58

Agricultural Marketing Service.

Yields of Corn in Michigan and Septemb~~er 1 Condition~~

The Agricultural Marketing Service estimates the probable production of many of the principal crops several months before they are harvested. Such estimates are based on returns obtained from farmers concerning the acreage planted and also on the farmers' judgment of "condition as a percent of normal." The reported condition of crops is one of the best available indications of the probable yield, assuming average weather conditions between the time of the report and the harvest of the crop.

The Division of Agricultural Estimates makes extensive use of dot charts and graphic analysis to interpret reported condition. One of the simplest types of graphic analysis used is that shown in the accompanying illustration. In this case we have plotted the September 1 condition and the final yield for each year from 1944 through 1953. The 10 observations lie fairly close to the curve we have drawn. Assuming that observations of the future will continue to cluster fairly closely around this curve, it could be used to estimate yields of corn in Michigan in future years.

In this case there may be some doubt about the curvature of the regression relationship. Conceivably, it could be a straight line rather than a curve which is concave upward. If we drew a straight line running approximately through the observations for 1949 and 1947, the deviations for the years 1952 and 1953 would be greater than from the curve we have drawn. However, this might reflect a net upward trend in yields of corn.

The way in which such charts are used by the Crop Reporting Board in making its estimates is described in detail in a publication entitled "The Agricultural Estimating and Reporting Services of the United States Department of Agriculture." 8/

8/ U. S. Dept. Agr. Mis. Pub. 703, 266 pp., illus. 1949.

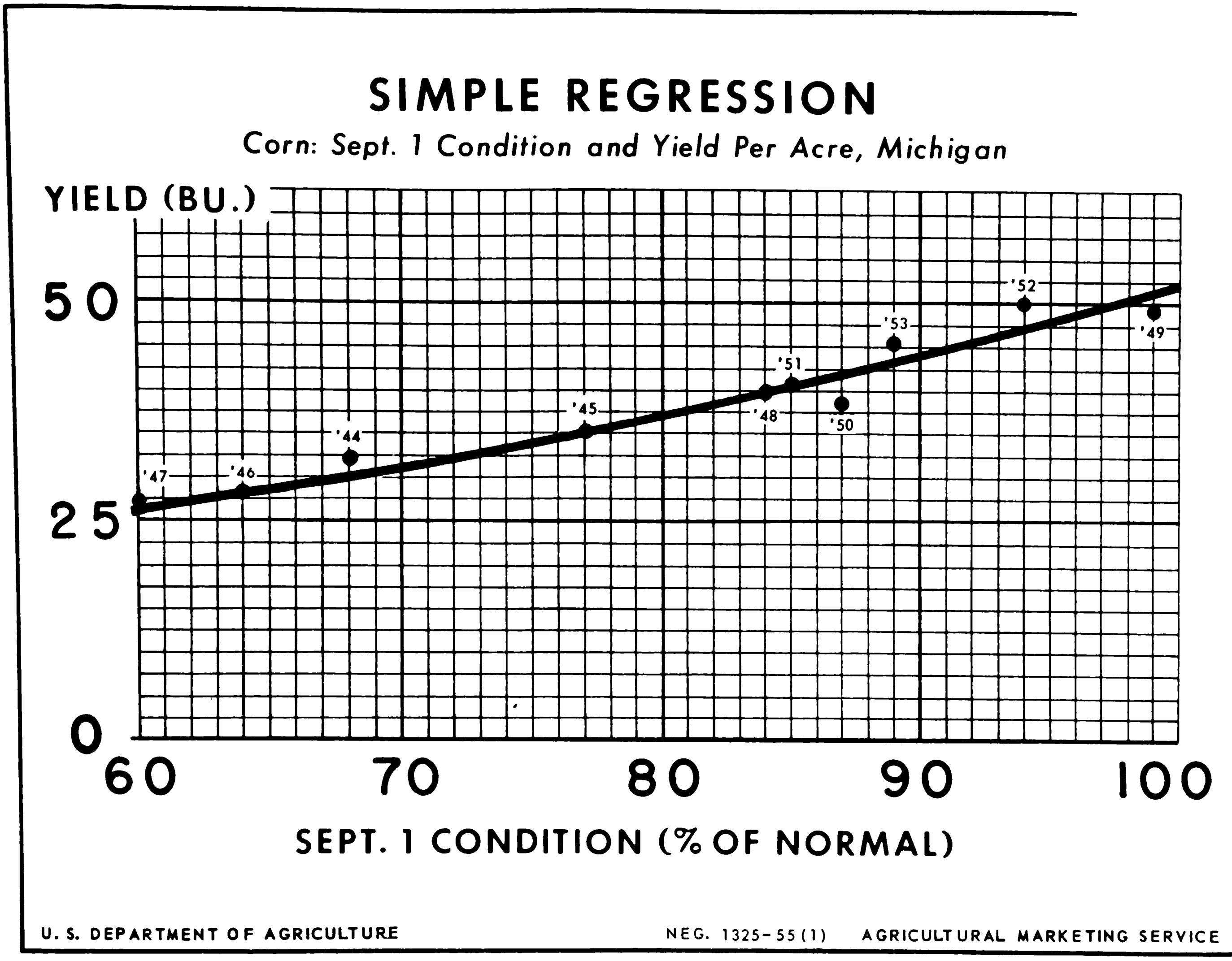


Figure 15

Corn: Condition September 1 and yield per acre, Michigan, 1944-53

Year	Condition 1/	Yield
	Percent	Bushels
1944	68	32.0
1945	77	35.0
1946	64	28.0
1947	60	27.5
1948	84	39.5
1949	99	49.0
1950	87	38.5
1951	85	41.5
1952	94	50.0
1953	89	45.5

1/ As a percentage of normal.

Agricultural Marketing Service.

Food Prices, Consumer Incomes, and Volume of Food Marketings

Economists often work with time series; that is, with records of prices, production, and consumption over a period of time. When studying relations between time series, particularly if several variables are involved, it is a good practice to plot each series before drawing dot charts such as the ones we have just discussed.

Suppose, for example, that we were trying to discover the factors which affect retail food prices. Two of the factors that would doubtless come to mind are consumer incomes and the volume of food marketings. Before rushing to the calculating machine or even drawing a dot chart, it would be a good idea to plot each series as we have done in this diagram and to study the changes which have occurred over a period of time.

In this case it is clear that there is high correlation between the food price index and per capita disposable income. In fact, the relationship is so pronounced that it tends to overshadow the effect of per capita food marketings. We might notice, too, that during the war years from 1941 to 1945 the relationships do not seem to be the same as in other years.

Comparisons of these three time series suggest that the correlation between the average price index for food and per capita disposable income would be reduced by deflating each series (for example, by dividing each of these by the consumer price index for all commodities). Such a computation would also reduce the magnitude of the gyrations to more nearly correspond to those for per capita marketings of food. The sharp rise in marketings of food during the war years and subsequent decline, which appears to have taken place independent of changes in the other series, suggests that the war years be omitted from the analysis. If a chart of this sort indicates pronounced trends in one or more variables, it suggests that the analysis might yield improved results if it were based on year-to-year changes in the variables.

In some cases, a comparison of time series will indicate a timelag between changes in one variable and changes in another. We saw previously that changes in slaughter of hogs occur several months after a change in the ratio of hog prices to those for corn.

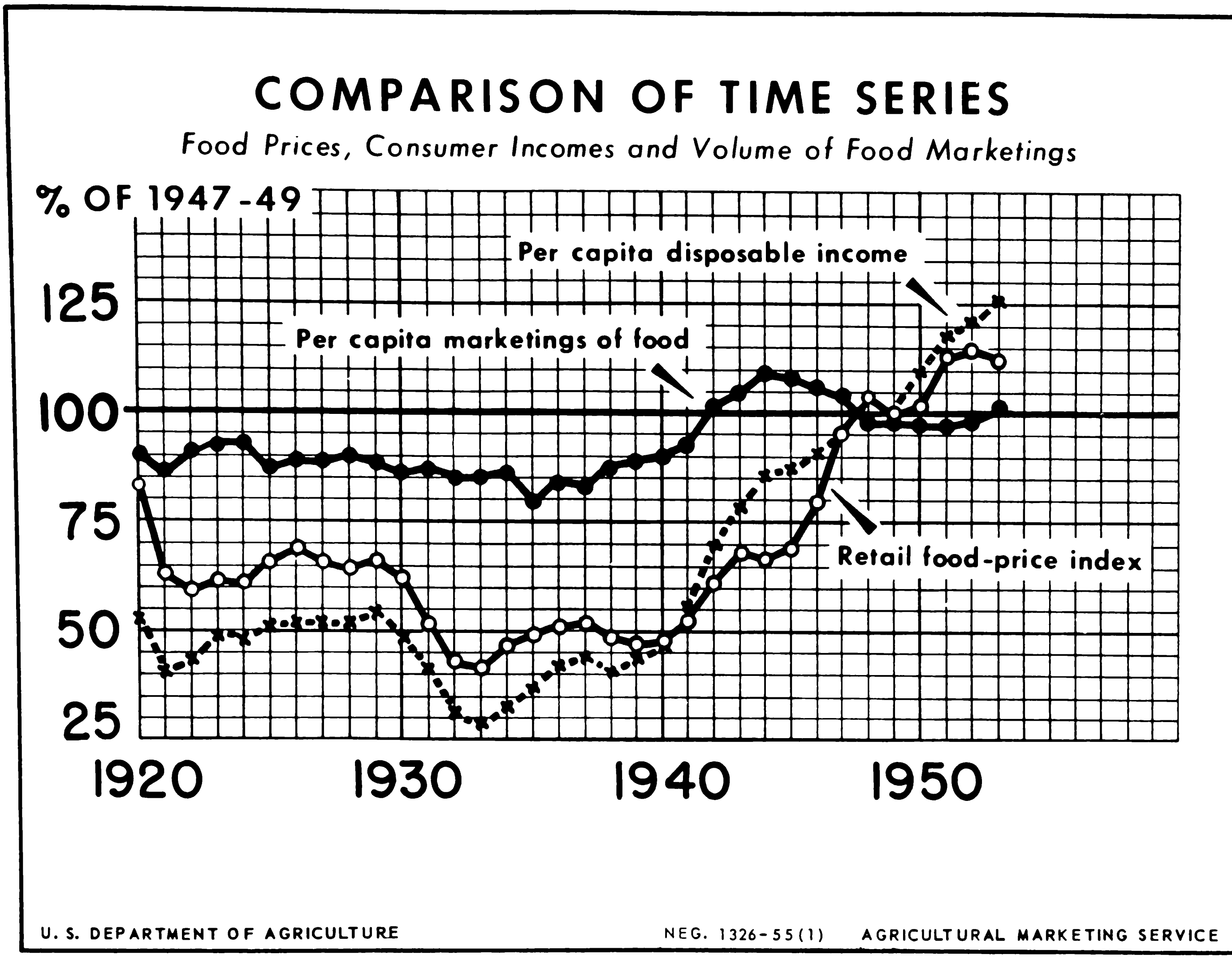


Figure 16

Price and marketing of food and disposable income: Index numbers, 1920-53

[1947-49=100]

Year	Retail price of food	Per capita: Marketing of food	Per capita: Disposable income	Year	Retail price of food	Per capita: Marketing of food	Per capita: Disposable income
1920	83.6	90	52.5	1937	52.1	83	44.5
1921	63.5	86	40.7	1938	48.4	87	40.7
1922	59.4	91	43.4	1939	47.1	89	43.3
1923	61.4	93	49.4	1940	47.8	90	46.2
1924	60.8	93	48.9	1941	52.2	93	55.6
1925	65.8	87	50.9	1942	61.3	101	69.9
1926	68.0	89	52.1	1943	68.3	105	78.2
1927	65.5	89	51.6	1944	67.4	109	85.7
1928	64.8	90	52.3	1945	68.9	108	87.2
1929	65.6	88	54.7	1946	79.0	106	90.7
1930	62.4	86	48.3	1947	95.9	104	95.0
1931	51.4	87	41.0	1948	104.1	98	103.7
1932	42.8	85	30.9	1949	100.0	98	101.3
1933	41.6	85	29.1	1950	101.2	97	109.5
1934	46.4	86	33.0	1951	112.6	97	117.7
1935	49.7	79	36.7	1952	114.6	98	120.8
1936	50.1	84	41.7	1953	112.5	101	125.4

Prices from Bureau of Labor Statistics, marketings of food from Agricultural Marketing Service, and income from Department of Commerce.

Price of Corn Related to Prices of Livestock and Livestock Products and Supply of Feed Concentrates Per Animal Unit

The graphics of simple (2-variable) regression is quick and easy. The graphics of multiple (several-variable) regression is naturally more complicated. But, as Bean 9/ showed, graphics can handle these problems, too. In multiple correlation we try to estimate the expected value of some dependent variable from given values of two or more independent variables. We assume an additive relation, such as:

$$X_o = f_1 (X_1) + f_2 (X_2)$$

and our problem is to estimate these functions.

The diagram illustrates an analysis of corn prices (X_0) related to two independent variables--prices of livestock and livestock products (X_1) and supplies of feed concentrates per animal unit (X_2). We know from theory and from general observation that high livestock prices tend to be associated with high prices of corn. We also know that large supplies of feed concentrates tend to be associated with low prices of corn. But we want to quantify these relationships--perhaps to forecast prices of corn.

Section A of this chart shows corn prices and prices of livestock and livestock products from 1936 through 1951. The dots are not clustered closely around any smooth curve--indicating that the simple (2-variable) correlation is rather low. Before drawing the regression line, we try to take account of X_2. We draw several regressions for subsamples of data, commonly called "drift lines." Thus in 1948, 1949, and 1950, supplies of concentrates were from 1.03 to 1.06 tons. We connect these observations with a drift line. Similarly we connect the observations for 1940, 1941, and 1942, when supplies were 0.90 tons. After drawing all possible drift lines, we draw a net regression line the slope of which represents approximately an average of the slopes of the drift lines. In this case, a straight line happens to be satisfactory. In many cases, a curve would be indicated.

Section B shows how the residuals (departures from the first regression line) are related to X_2. These residuals are clustered closely around the regression line we have drawn. If a nearly perfect fit were not given by the dots around this line, the process of successive approximation would be used. Foote 10/ has shown that when we use this method graphically based on linear relationships, the slopes of the successive approximations tend to converge toward the value that would be obtained had we fitted a mathematical regression line by the method of least squares.

Graphic multiple regression requires a fair amount of imagination and some practice. But it often shows up important relationships that are not brought to light by grinding figures out of a computing machine.

9/ Bean, Louis H. A Simplified Method of Graphic Curvilinear Correlation. Jour. Amer. Statis. Assoc. 24:386-397, illus. 1929.

10/ Foote, Richard J. The Mathematical Basis for the Bean Method of Graphic Multiple Correlation. Jour. Amer. Statis. Assoc. 48:778-788. 1953.

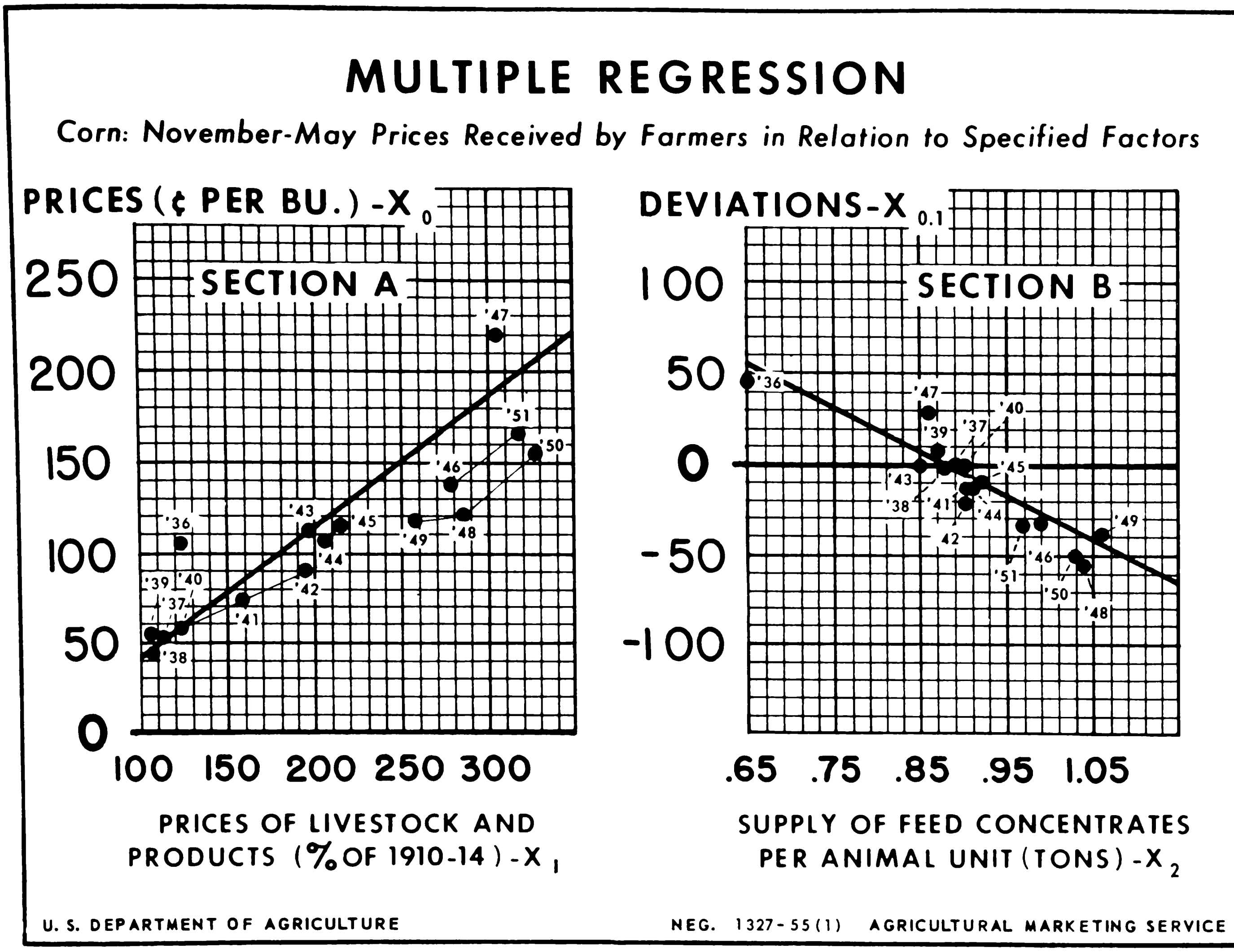


Figure 17

Corn: Price per bushel received by farmers and related variables, 1936-51

Period beginning	Price received by farmers (Nov.-May): Corn	Price received by farmers (Nov.-May): Livestock and products 1/	Supply of feed concentrates per animal unit 2/
	Cents		Tons
1936	106	123	0.65
1937	51	114	.89
1938	44	108	.88
1939	55	107	.87
1940	58	122	.90
1941	74	159	.90
1942	90	194	.90
1943	112	196	.85
1944	107	206	.91
1945	115	215	.92
1946	138	278	.99
1947	220	305	.86
1948	120	285	1.04
1949	118	258	1.06
1950	155	327	1.03
1951	167	318	.97

1/ Index number, 1910-14=100. 2/ Year beginning October.

Computed from data in Foote, Richard J. Statistical Analyses Relating to the Feed-Livestock Economy. U. S. Dept. Agr. Tech. Bul. 1070. 1953. p. 6.

Production of California Bartlett Pears and August 1 Condition

The condition of California Bartlett pears as reported by farmers on August 1 gives a good indication of probable production. However, in this case use of a multiple regression analysis helps to refine the relationship. A reported condition of 80 percent in recent years indicates a production considerably higher than we would have expected from the same condition figure 5 or 10 years ago.

Section A of the diagram is a simple dot chart showing the relation of August 1 condition to production. As in the case of several other diagrams, we have drawn some light lines covering the observations in chronological order. Then we have drawn the indicated net regression line.

Section B of the diagram shows the net trend in production. The residual (that is, the difference between actual and estimated production in Section A) was plotted for each year in succession. Then a line that fits the points approximately was drawn to indicate the net rate of increase over the 10-year period.

The production of California Bartlett pears, like that of other fruit crops, depends upon many things including the number and ages of trees, changes in production techniques, and similar factors. Many of these factors are not fully reflected in the reported condition. It is, therefore, essential when dealing with certain crops to consider whether factors not measured by condition have changed consistently over time. If so, a net trend, as drawn here, measures the increase or decrease that we would expect after allowing for the effects of changes in condition throughout the period.

In handling problems of this type, a number of approaches can be used. For example, deviations from the line or curve in section A may be expressed as percentages of production before using them to measure the trend in section B.

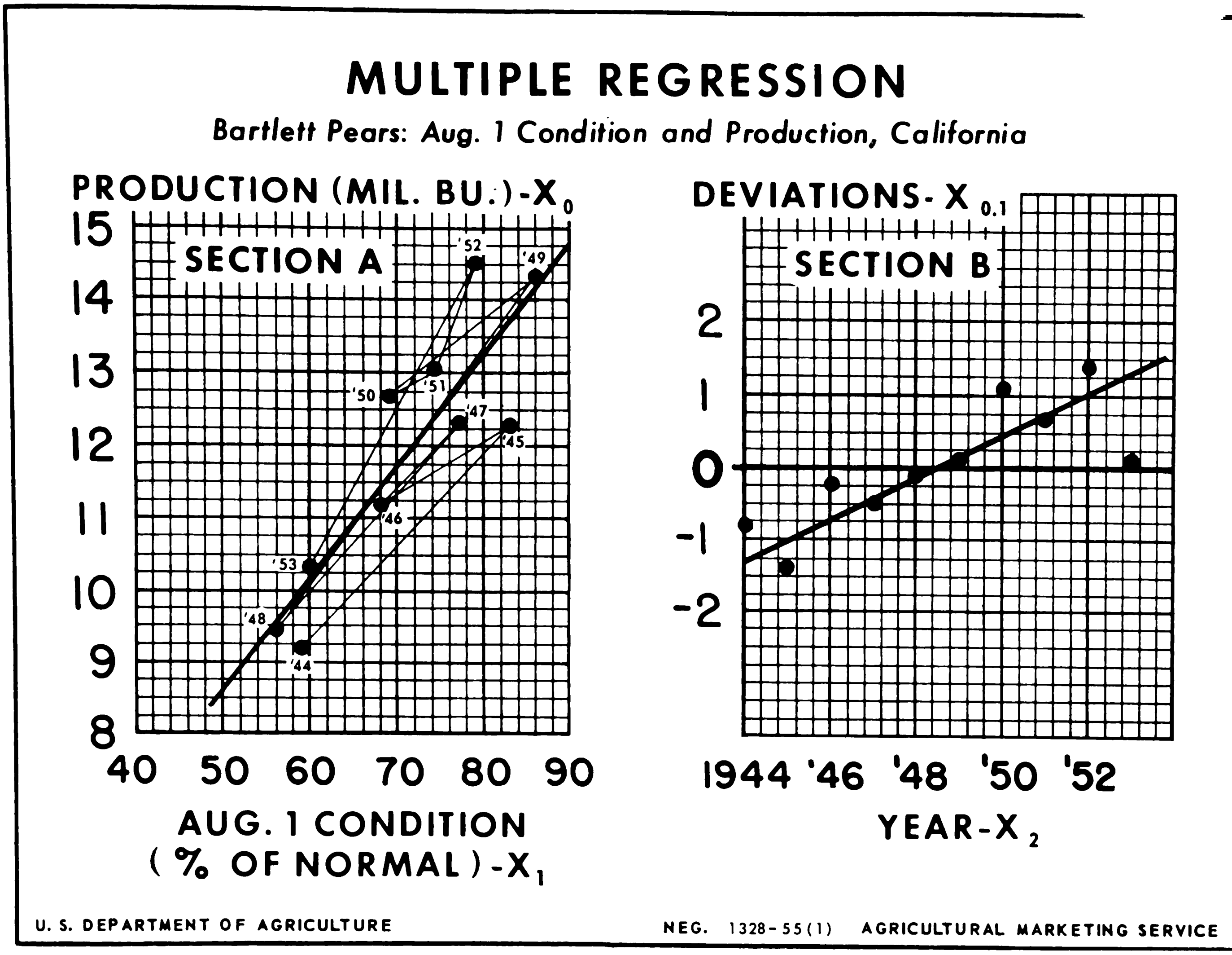


Figure 18

Bartlett pears: Condition August 1 and production, California, 1944-53

Year	Condition 1/	Production
	Percent	Million bushels
1944	59	9.2
1945	83	12.3
1946	68	11.2
1947	77	12.3
1948	56	9.4
1949	86	14.3
1950	69	12.7
1951	74	13.0
1952	79	14.5
1953	60	10.3

1/ Percentage of normal.

Agricultural Marketing Service.

Price of Late Onions Related to Production and Disposable Income

The data for this diagram, taken from Shuffett, 11/ are expressed as first differences (i.e. year-to-year changes) in logarithms. The rationale of this may be found in Shuffett's bulletin and need not concern us here. Graphic analysis will handle logarithms and first differences, as well as the unmanipulated data.

The main purpose of this diagram is to illustrate successive approximations to the true regression lines. We have already discussed the graphic determination of the net regression lines. So far, we have tacitly assumed that one approximation is enough. But in many cases the statistician should try two or more successive approximations.

The original data (here they are the first differences of logarithms) are plotted as in the regression charts we have already discussed. The black dots in section A show the joint scatter of production and price. The heavy line is our first approximation to the net regression of production on price. (Drift lines were drawn, but have been erased to keep from cluttering up the chart.) Deviations from this line were then plotted as heavy dots in section B. The solid line through these heavy dots is the first approximation of the net regression of disposable income on price.

So far, our analysis is the same as in several previous diagrams. We now proceed to make a second approximation. The deviations from the solid line in section B are now plotted as circles in section A. The dashed line, drawn through these circles, is our second approximation to the net regression of production on price. Then the deviations from this dashed line are plotted as circles in section B. A dashed line, drawn to fit these circles, is our second approximation to the net regression of disposable income on price.

This process can be continued to get as many approximations as needed. If done correctly, the successive approximations will converge to the true (least squares) regressions. Ordinarily two or three approximations are enough.

11/ Shuffett, D. Milton. The Demand and Price Structure for Selected Vegetables. U. S. Dept. Agr. Tech. Bul. 1105, pp. 38-43. 1954.

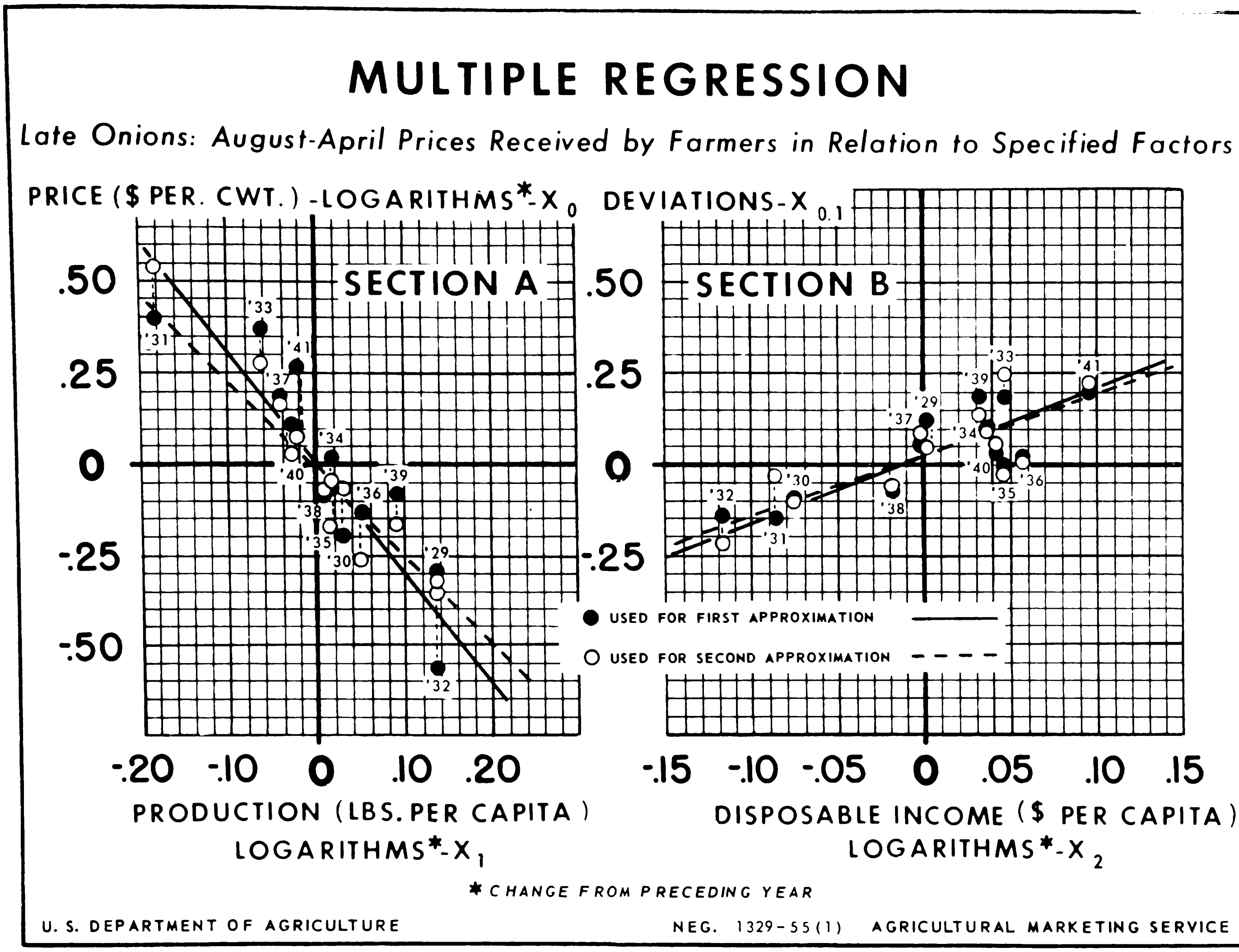


Figure 19

Late onions: Average price per 100 pounds received by farmers and related variables, August-April 1928-41

Period beginning	Price 1/ Actual	Price 1/ First difference of: logarithms	Per capita Production 1/ Actual 2/	Per capita Production 1/ First difference of: logarithms	Per capita Disposable income Actual 3/	Per capita Disposable income First difference of: logarithms	Period beginning	Price 1/ Actual	Price 1/ First difference of: logarithms	Per capita Production 1/ Actual 2/	Per capita Production 1/ First difference of: logarithms	Per capita Disposable income Actual 3/	Per capita Disposable income First difference of: logarithms
	Dollars		Pounds		Dollars			Dollars		Pounds		Dollars	
1928	2.54	---	6.65	---	658	---	1935	1.18	-.06166	8.20	.01673	467	.04607
1929	1.30	-.29089	9.08	.13527	663	.00328	1936	.86	-.13738	9.23	.05139	534	.05822
1930	.82	-.20013	9.75	.03091	557	-.07565	1937	1.30	.17944	8.36	-.04299	532	-.00163
1931	2.02	.39154	6.41	-.18214	456	-.08690	1938	1.06	-.08863	8.59	.01178	509	-.01919
1932	.54	-.57296	8.75	.13515	347	-.11863	1939	.88	-.08083	10.57	.09009	546	.03047
1933	1.28	.37482	7.58	-.06234	386	.04626	1940	1.12	.10474	9.93	-.02713	601	.04168
1934	1.36	.02633	7.89	.01741	420	.03666	1941	2.08	.26884	9.47	-.02060	748	.09503

1/ Excludes quantities produced in market gardens for sale in nearby cities prior to 1939.
2/ Production divided by November 1 civilian population.
3/ Disposable income at annual rates divided by November 1 civilian population.

Shuffett, D. Milton. The Demand and Price Structure for Selected Vegetables. U. S. Dept. Agr. Tech. Bul. 1105. 1954. p.43.

Ascorbic Acid in Snap Beans Related to Storage Time and Temperature

Sometimes the graphic method of multiple correlation is criticized as being too flexible. Actually, it is not flexible enough to describe at all accurately some of the common relationships in physical and economic science. Any linear multiple regression, whether it is determined by mathematical computations or by graphics, assumes that the effects of two or more independent variables can be added together to estimate the dependent variable. Joint regression assumes a more general relationship between the variables. In the 3-variable case, we need to determine a smooth 3-dimensional surface describing how one variable changes in relation to two others.

One technique for doing this is similar to that used in surveying and grading land. We can forget for the moment that the chart refers to snap beans. Suppose that the vertical axis measured distances north and south, the horizontal axis measured distances east and west, and the numbers written by the dots on the diagram indicated the elevation of the land at various points as determined by a surveyor's transit. Anyone used to maps would recognize that the land is level at the left side of the diagram and that it slopes rather steeply at the right side, going down hill as we go up on the diagram. He would also see that there are bumps and hollows. If a landscape gardener were going to improve this plot of land, he would smooth out the surface. In simple regression we smooth in only one dimension. Here we are smoothing in two dimensions. We can describe the general lay of the land by a series of smooth contour lines.

Of course, we are not dealing here with land and contour maps. However, the general problem of joint regression is that of determining a series of isoquants. Whatever the three variables may be, an isoquant will show the combinations of two independent variables which correspond to a given value of the dependent variable. In the case illustrated by the diagram, it is clear that storage time has little or no effect upon ascorbic acid in snap beans if they are held at a temperature of 0° F. Regardless of length of storage, the beans contain nearly 16 percent of acid. As the temperature increases above 0, however, storage affects ascorbic acid more and more. At 20° F, for example, snap beans lose ascorbic acid very rapidly. The isoquant labeled 13 indicates that the beans will contain roughly 13 milligrams of acid per 100 grams with storage for a very short time at 20 degrees, with storage of 4 weeks at about 14 degrees, with storage of 6 weeks at 11 degrees, or with storage of 8 weeks at 8 degrees.

With a little practice anyone can draw isoquants graphically, as we have in this diagram, that give at least a general indication of the relationships involved. If the researcher wants to fit mathematical functions, the diagram should suggest the kind of function to use. In this case, the series of isoquants look something like a spiral staircase, with the stairs becomming steeper as they go down. The formula for a spiral staircase is simple enough mathematically if anyone cared to fit it.

Another technique which is sometimes used to study three-dimensional relationships is the "isometric projection." Those who are not familiar with isoquants may find such projections easier to visualize. But they are also harder to read accurately.

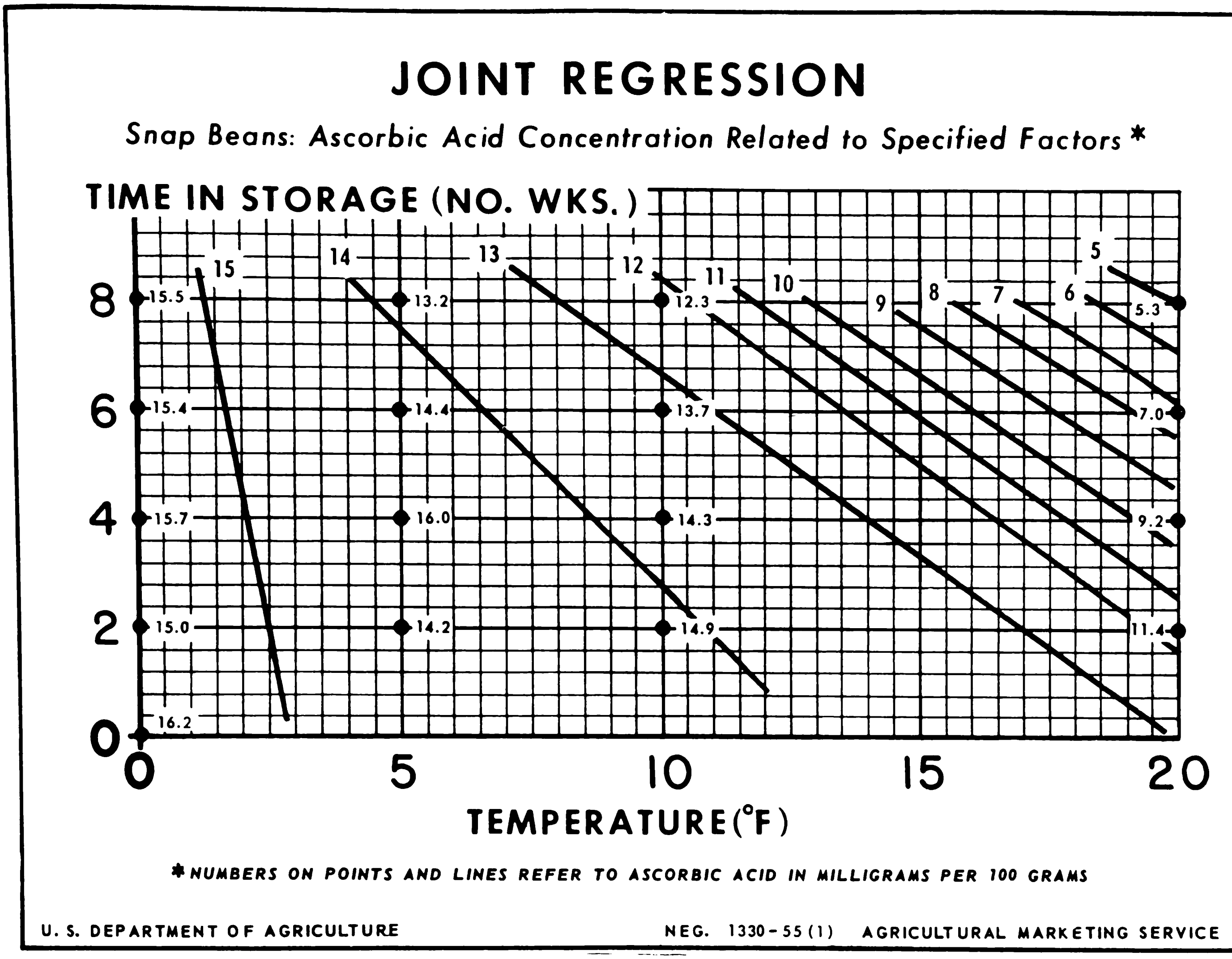


Figure 20

Snapbeans: Average concentration of ascorbic acid per 100 grams, by temperature and time in storage

Time in storage	Concentration with storage temperature in degrees Fahrenheit			
	0	5	10	20
Weeks	Milligrams	Milligrams	Milligrams	Milligrams
0	16.2	---	---	---
2	15.0	14.2	14.9	11.4
4	15.7	16.0	14.3	9.2
6	15.4	14.4	13.7	7.0
8	15.5	13.2	12.3	5.3

Snedecor, George W. Application of the Theory of Experimental Design in Biology. Proc. of Int. Statis. Inst. 3:440-52. 1947.

Combination of Two Farm Enterprises

Programming is the planning of economic activities to maximize income or to minimize costs. In some cases it is reasonable to assume that the input-output relationships are approximately linear. For example, if we know how much seed, labor, and fertilizer is required to grow an acre of potatoes, we can assume that it will require about twice as much of each input factor to grow two acres of potatoes by the same process. In a similar manner, if we know the amount of protein, calcium, and other nutrients in a bushel of corn, there would be twice as much of each nutrient in two bushels of corn. These are linear relationships, and in cases of this kind we can estimate the optimum program by a technique known as linear programming.

The data on this chart show two possible farm enterprises in North Carolina and six input factors. 12/ To be feasible a combination of inputs must not require more than the available amount of any resource. In an analysis of this kind it is convenient first to compute for each enterprise the proportion of available resources needed to produce some arbitrary amount of net income. In this case we chose $10,000. For example, to get a net income of $10,000 from beef cattle would require 4.63 times as much spring land as the farmer has available. The left scale of the chart represents the proportions of available resources needed to get a net income of $10,000 from beef cattle. The right scale shows the proportion of available resources needed to produce $10,000 of net income from fall cabbage. If we had to choose one or the other of these enterprises, the choice should be fall cabbage, since the highest dot on the right scale is lower than the highest dot on the left scale. The limiting factor for fall cabbage is September-October labor. To get an income of $10,000 from fall cabbage would require 2.17 times as much September-October labor as the farmer has available. If he used all of his September-October labor on cabbage, his income would be $10,000 divided by 2.17, or $4,608. This is better than he could get from beef cattle alone.

However, this farmer could raise his income by combining beef cattle with fall cabbage. Each of the six lines drawn across the diagram show the proportion of some resource needed for various combinations of beef cattle and fall cabbage. The limiting factor for any combination is indicated by the top line at that point on the horizontal scale. A combination that is mostly beef cattle has as its limiting factor fall land. With combinations including 46 to 91 percent fall cabbage, the limiting factor is production capital. Finally, in combinations that are mostly fall cabbage and only a little beef cattle, the limiting factor is September-October labor. The minimax point (that is, the lowest of the maximum points for any combination) indicates that the most profitable combination of these two enterprises would use about 91 percent of (1) the available production capital and (2) the September-October labor to produce fall cabbage. The other 9 percent of these two limiting factors would be used for beef cattle. To get an income of $10,000 from these combinations would require almost twice as much of the two factors as are available. So the best the farmer could get with these two enterprises would be an income of a little over $5,000.

12/ See King, R. A., and Freund, R. J. A Procedure for Solving a Linear Programming Problem. N. C. Agr. Expt. Sta. Jour. Paper 503, 18 pp. 1953. (Processed.) This study lists 9 different inputs needed to carry on each of 6 different enterprises.

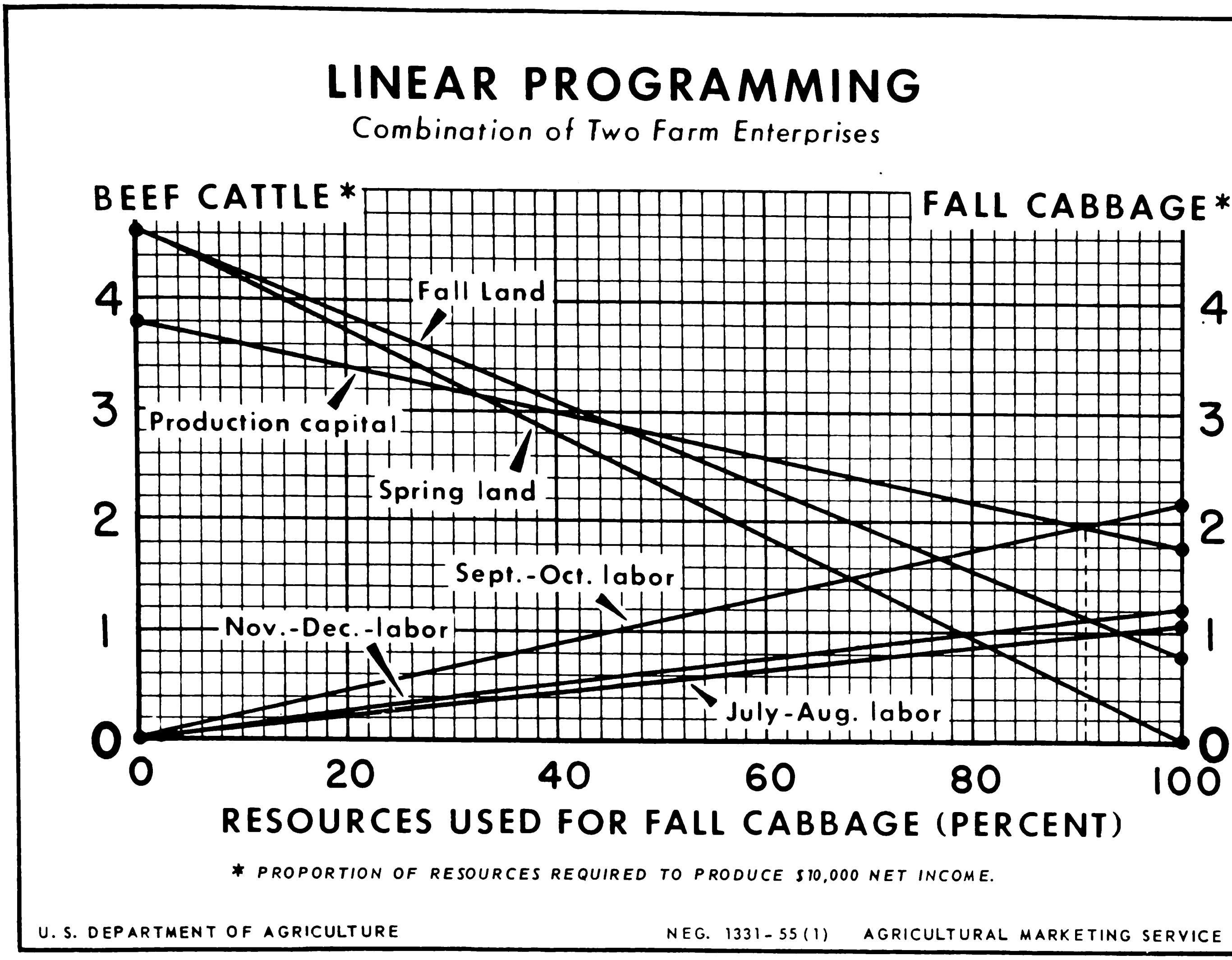


Figure 21

Beef cattle and fall cabbage: Proportion of available resources required to produce $10,000 net income for a farm, North Carolina

Resource	Proportion required	
	Beef cattle	Fall cabbage
Land:		
Spring	4.63	0
Fall	4.63	0.80
Production capital	3.78	1.80
Labor:		
July-August	0	1.08
September-October	0	2.17
November-December	0	1.22

King, R. A. and Freund, R. J. A Procedure for Solving a Linear Programming Problem. N. Ca. Agr. Expt. Sta. Jour. Paper 503. 1953. (Processed.) p. 13.

The Minimum-Cost Dairy Feed

Here is another diagram that is useful in linear programming. In this case, we want the least cost combination of feeds that will meet stated requirements. The prices of several feeds are given; also such requirements as total digestible nutrients and protein. We first compute the proportion of each requirement that could be supplied by $1 worth of corn, $1 worth of oats, etc. The net result is shown on the table and plotted on the chart.

We then consider combinations of two feeds that will meet two requirements--those for total digestible nutrients and for protein. For $1 we could buy any combination lying along a straight line joining two dots. We have drawn such a line showing combinations of gluten and middlings. A balanced ration would lie on a line through the origin having a slope of 45 degrees. The point at which this line cuts the line connecting the points for gluten and middlings indicates a ration mostly of gluten with a small amount of middlings. It can be shown that this combination will meet the two requirements at less expense than either feed alone. This is true because (1) the line joining the two dots slopes downward to the right and (2) it crosses the 45-degree line. If these two conditions were not met, it would be less expensive to meet the two nutritive requirements from a single feed. Also, this combination is less expensive than any other combination of two feeds that would meet the two nutritive conditions. This is because no dot lies above the line (extended by dashes) joining the dots for gluten and middlings. If there were a dot above this line it would indicate that the cost would be reduced by substituting this feed for one of those in the combination. If the combination of gluten and middlings not only meets the requirements for total digestible nutrients and for protein, but also meets all other requirements, the combination we have found is the final answer--that is, it will meet all requirements at less expense than any other possible combination of feeds. This example is discussed in more detail in an article published in 1951. 13/

In each of these cases, we have shown only combinations of two enterprises. In linear programming we need to study other pairs. This can be done quickly by the graphic method. This method is more difficult when we consider combinations of three enterprises, and becomes impossible when we consider more than three. In such cases we need to use the so-called "simplex technique." 14/ However, even when we use the simplex technique, the first two or three steps should be done graphically. The diagrams shown here are similar to those used by Dorfman. 15/

13/ Waugh, Frederick V. The Minimum-Cost Dairy Feed. Jour. Farm Econ. 33:299-310, illus. 1951.

14/ This is described in Dantzig, George B., Maximization of a Linear Function of Variables Subject to Linear Inequalities and Application of the Simplex Method to a Transportation Problem, and Dorfman, Robert, Application of the Simplex Method to a Game Theory Problem. In Koopmans, Tjalling C., ed. Activity Analysis of Production and Allocation. Cowles Commission for Research in Economics Monogr. 13, pp. 339-373. New York. 1951.

15/ Dorfman, Robert. Mathematical, or "Linear," Programming: A Nonmathematical Exposition. Amer. Econ. Rev. 43:797-825, illus. 1953.

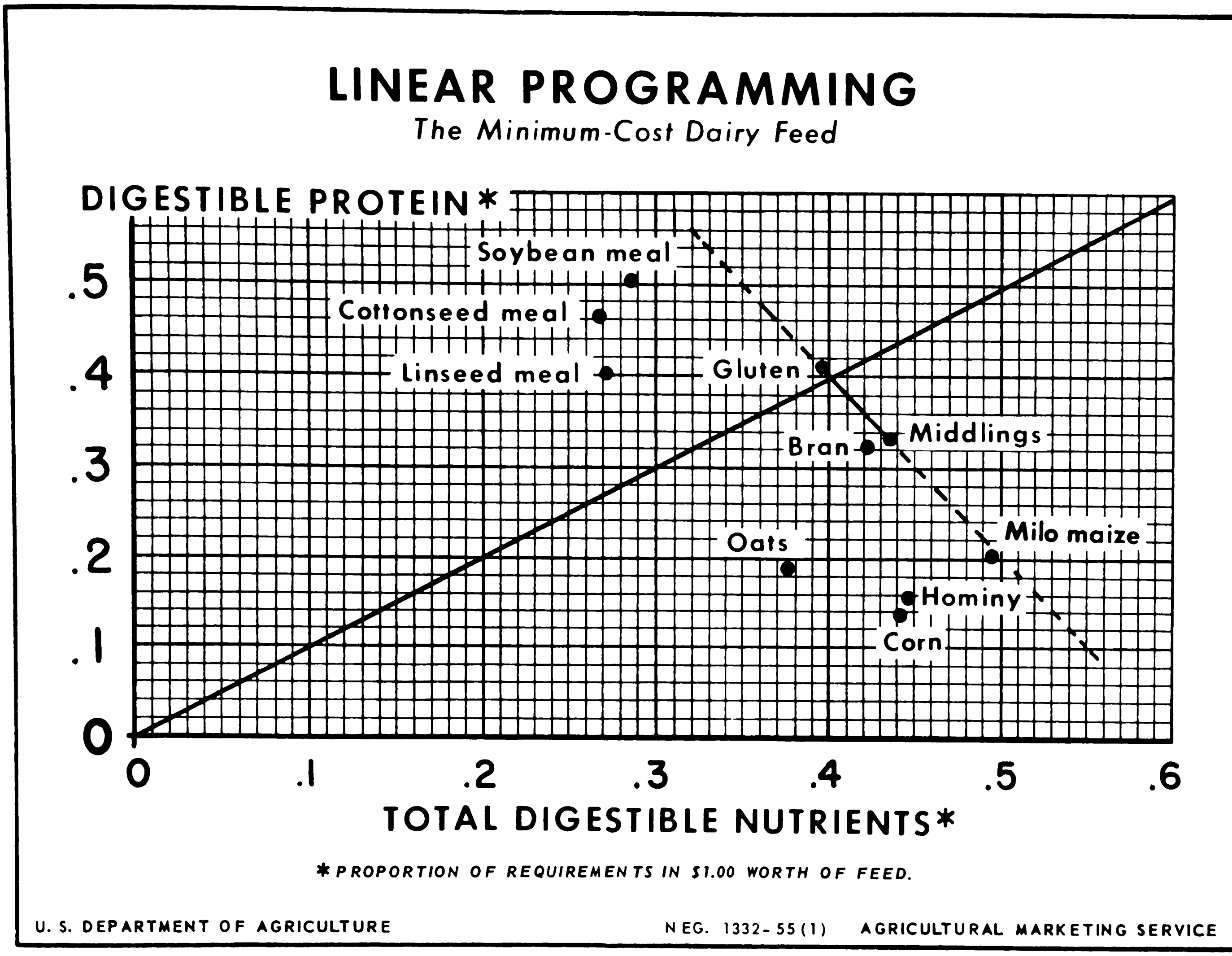


Figure 22

Dairy feed: Proportion of the requirements for protein and total digestible nutrients supplied by $1 worth of each feed

Feed	Proportion supplied: Digestible protein	Proportion supplied: Total digestible nutrients
Corn	0.136	0.441
Oats	.187	.375
Milo maize	.203	.495
Bran	.321	.423
Middlings	.332	.436
Linseed meal	.400	.272
Cottonseed meal	.464	.268
Soybean meal	.504	.286
Gluten	.412	.395
Hominy	.158	.448

Waugh, Frederick V. The Minimum Cost Dairy Feed. Journal Farm Economics. 33: 299-307, illus. 1951.

Beef and Pork

Economists have often discussed the possibility (or the impossibility) of deriving a set of indifference curves from market data. But they have seldom tried it. The pure economic theory of indifference curves is important. But in the practical analysis of such problems as cost-of-living indexes and the incidence of taxes, we need reasonably accurate lines or curves derived from market data. In some cases, at least, I believe that we can obtain approximate indifference functions by a simple graphic method. This is particularly true if the income elasticities for two goods are approximately equal. I hope to discuss this problem in more detail elsewhere and to justify the method illustrated by this example.

In each of the sections of the accompanying diagram, each x (which looks like a short cross-line) indicates the combination of beef and pork bought per capita in one year. The heavy line drawn through the x shows combinations that could have been bought with a given expenditure. Take 1932, for example. (See section B.) A typical consumer, buying the average per capita amounts, purchased 46.0 pounds of beef and 69.7 pounds of pork. The price of beef was 1.596 times the price of pork. So with the same expenditure he could have bought 1 less pound of beef and 1.596 more pounds of pork--or 10 pounds less of beef and 15.96 more of pork. The fact that he chose to buy 46.0 pounds of beef and 69.7 pounds of pork shows that he preferred this combination to the others on the heavy straight line. An indifference curve must, therefore, be tangent to the heavy straight line at the point marked with an x, and, similarly, with the other lines and points on the diagram. Also, we know that no two indifference curves can cross one another. Our problem, then, is to draw a set of curves in each section meeting two conditions: (1) Each curve must be (approximately) tangent to the heavy straight line at the point marked x, and (2) no pair of curves can cross one another.

The family of curves shown on the diagrams were drawn graphically following the two rules stated above. The fit is excellent in almost all years. It is not perfect, as is the case of most statistical work. The slopes and curvatures <u>must</u> be approximately as we have drawn them. Otherwise they would conflict either with the statistics or with the logical conditions that must be met. We assume that the consumer prefers larger amounts of meat to smaller amounts. Thus, the least preferred combinations are those nearest to the lower left part of each diagram; the most preferred are at the upper right.

The analyses based on data for 1921-31 and 1932-41 and a similar one based on 1948-53 indicate that important shifts have taken place in the relative demand for these items over time. Within any of the three periods, indifference curves can be drawn that are approximately tangent to the indifference lines at the point for which an observation is available and that meet the theoretical requirements. However, if the data from the three periods are combined in a single chart, such lines cannot be drawn. A research project currently underway in the Agricultural Marketing Service will attempt to explain why this is so.

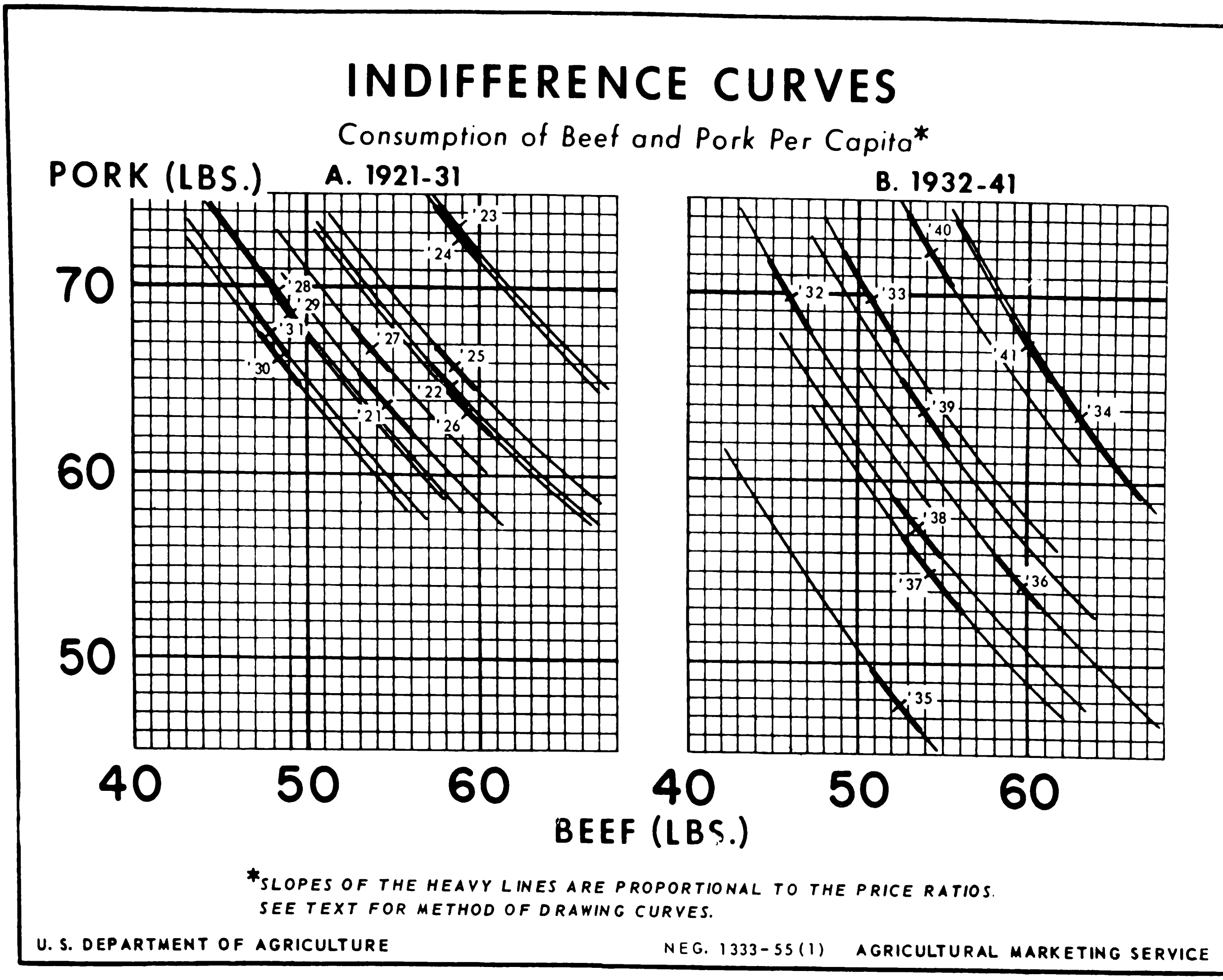


Figure 23

Beef and pork: Ratio of beef price to pork price at retail and per capita consumption, 1921-41

Year	Retail price per pound		Price ratio	Consumption		Year	Retail price per pound		Price ratio	Consumption	
	Beef	Pork		Beef	Pork		Beef	Pork		Beef	Pork
	Cents	Cents		Pounds	Pounds		Cents	Cents		Pounds	Pounds
1921	29.3	28.1	1.043	54.7	63.9	1932	24.9	15.6	1.596	46.0	69.7
1922	27.7	26.8	1.034	58.3	64.8	1933	21.5	13.9	1.546	50.8	69.8
1923	28.8	25.3	1.138	58.8	73.2	1934	23.3	18.8	1.239	63.0	63.6
1924	29.5	25.3	1.166	58.7	73.0	1935	30.5	27.4	1.113	52.5	47.7
1925	30.7	31.1	.987	58.6	65.8	1936	28.6	26.9	1.063	59.7	54.4
1926	31.4	33.3	.943	59.4	63.3	1937	32.5	27.7	1.173	54.4	55.0
1927	32.8	31.2	1.051	53.7	66.8	1938	28.7	24.5	1.171	53.6	57.4
1928	37.4	29.5	1.268	48.1	69.9	1939	29.5	22.2	1.329	53.9	63.9
1929	39.2	30.3	1.294	49.0	68.7	1940	29.5	19.3	1.528	54.2	72.4
1930	36.2	29.1	1.244	48.2	66.1	1941	31.5	24.7	1.275	60.0	67.4
1931	30.0	23.7	1.266	47.9	67.4						

Agricultural Marketing Service.

Gross Profit from Storage

In economic analysis we often want to compute the average of two or more points on a curve.

In this diagram the curve represents total returns to growers from sales of various amounts of eggs. In deriving these figures, allowance was made for the effect of disposable income on prices of eggs. The prices shown are those that might have been expected with income at its average level for the period 1940-48. A practical question is whether it would be profitable to store up the surplus in periods of large production and to sell it in periods of small production.

Suppose we produced 40 billion eggs in one period and 50 billion in another period. The returns for each period would be shown on the curve. The average for the two periods would be halfway between these two points. This average is indicated by the dot at the midpoint on the straight line joining the appropriate points on the curve. In this case it indicates a moderate gross profit from storage. That is, the gross income from selling 45 billion eggs in each period would be greater than the average income from selling 40 billion in the first period and 50 billion in the second. Costs of storage, handling, and any loss in quality would have to be deducted in order to determine whether net returns would be larger from storage.

It is easy to see that there will be a gross profit from storage if, and only if, the returns curve is concave downward. The degree of curvature is an important indication of the possible amount of gross profit.

Of course, this is only one of the many uses of averages. The economist-statistician often wants to compute average prices, average cost, average yield of a crop, and so on. When working with graphic diagrams, such averages can be computed graphically with little time or trouble. There is no need to read the numbers from the diagram, copy them on a piece of paper, add them, divide by two, and put the average back on the diagram. The simple arithmetic average of any two points on any curve can be located graphically by the graphic method explained here.

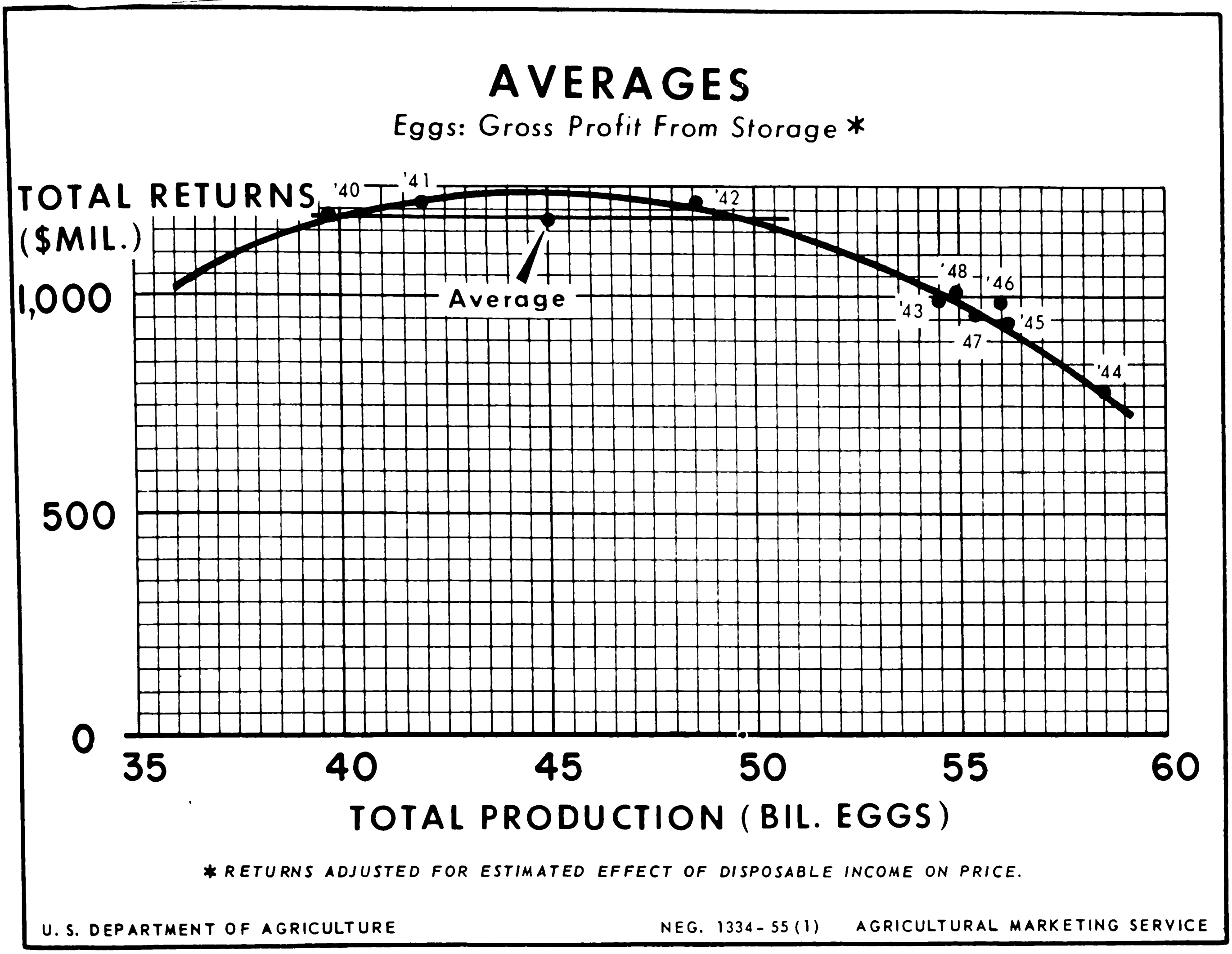


Figure 24

Eggs: Production, price per dozen received by farmers, and total returns, 1940-48

Year	Production	Price 1/	Total returns 1/
	Billions	Cents	Million dollars
1940	39.7	36	1,188
1941	41.9	35	1,225
1942	48.6	30	1,230
1943	54.5	22	990
1944	58.5	16	784
1945	56.2	20	940
1946	56.0	21	987
1947	55.4	21	966
1948	54.9	22	1,012

1/ Adjusted for estimated effect of disposable income on price.
Data derived from Figure 92 in Thomsen, Frederick L., and Foote, Richard J. Agricultural Prices. New York. 1952. p. 431.

Coefficient of Elasticity of Demand

Many economists have trouble with coefficients of elasticity. They are frequently concerned with the elasticity of demand--more precisely, with the elasticity of consumption with respect to price. The diagram shows how this can be measured graphically.

The curved line on the diagram represents an assumed demand curve for eggs. The scales for consumption and prices would not need to be shown. They are unimportant, because the coefficient of elasticity is invariant to changes in scale provided that the axes start at the origin. Suppose we want the coefficient of elasticity at the point ($p=a$, $q=c$). We draw the indicated straight line tangent to the demand curve at that point. The elasticity in question is $-a/b$. For this example, this equals -35.5 divided by 64.5 based on the scales shown. In terms of small squares on the grid, this equals 17.75 divided by 32.25. Either computation indicates an elasticity of -0.55.

This piece of graphics comes from Alfred Marshall. 16/ It derives from the definition of elasticity $\eta = \frac{dq}{dp} \cdot \frac{p}{q}$. Note that $\frac{dq}{dp} = -\frac{c+d}{a+b}$, and (by similar triangles) $\frac{c+d}{a+b} = \frac{c}{b}$. Also $p=a$, and $q=c$. So $\frac{dq}{dp} \cdot \frac{p}{q} = \frac{-c}{b} \cdot \frac{a}{c} = -a/b$.

Some economists have found the concept of elasticity so difficult that they have used "arc elasticity," or the "average elasticity of a curve." If the graphic approach to elasticity is used, there is little need for such concepts. The elasticity coefficient shown here is exact and easy to compute.

We should note that the concept of elasticity applies not only to demand curves--but to <u>any</u> curve. When we speak of the elasticity of demand we (usually) mean the elasticity of consumption with respect to price. But we might want the elasticity of cost of producing potatoes with respect to the amount of fertilizer used, for example. Whatever the curve, we can measure its elasticity at any point, using the same graphics as shown here.

16/ Marshall, Alfred. Principles of Economics. Ed. 8, pp. 102-103. New York. 1948. First published 1920.

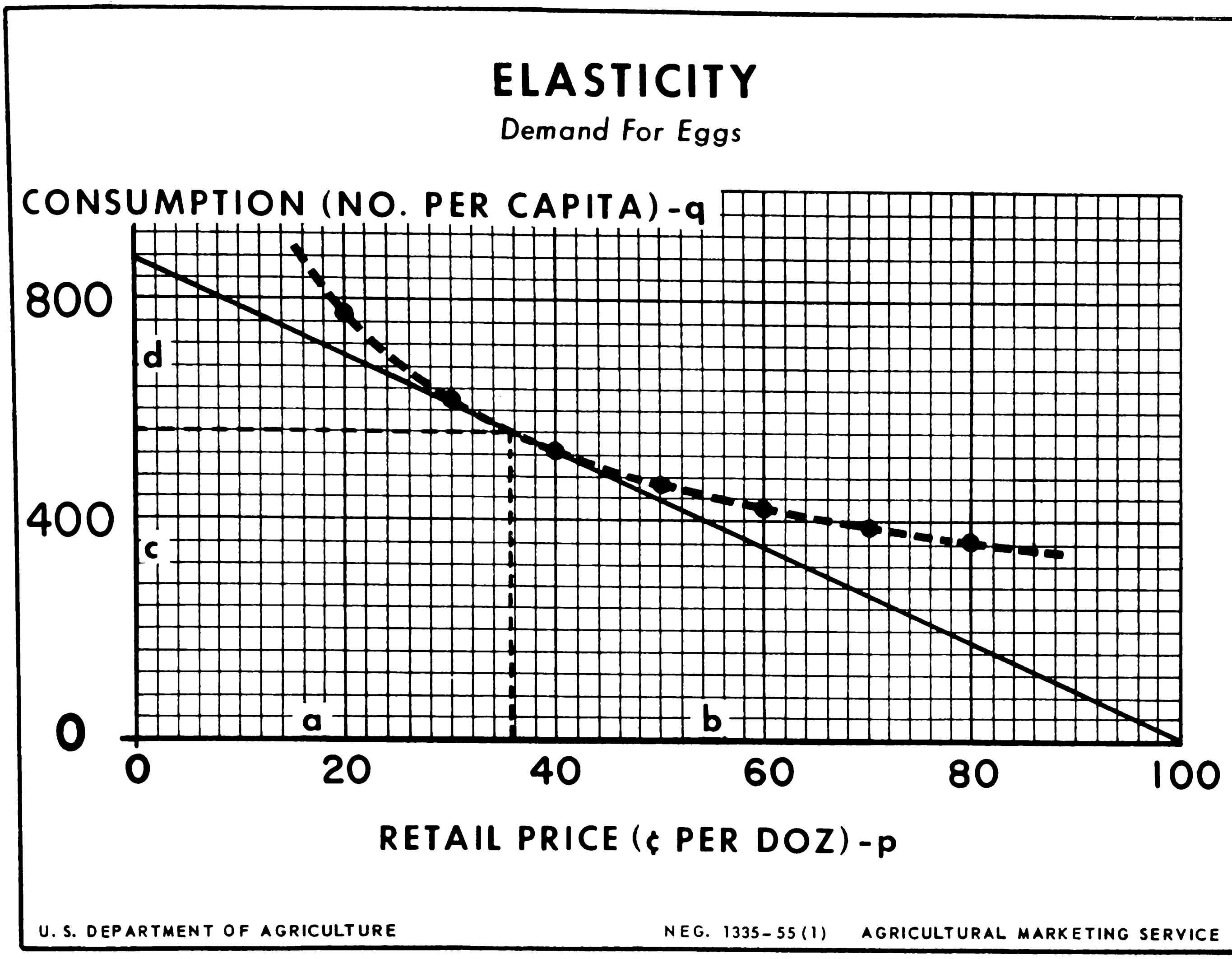


Figure 25

Eggs: Consumption per capita associated with given retail price per dozen

Consumption	Price
Number	Cents
777	20
621	30
530	40
469	50
424	60
390	70
362	80

Based on an assumed elasticity of demand coefficient of -0.55. See Foote, Richard J. and Fox, Karl A. Analytical Tools for Measuring Demand. U. S. Dept. Agr. Agr. Handbook 64. 1954. p. 40.

Some statisticians and economists find calculus a difficult subject. Differential calculus is relatively easy if you do it graphically. The differential at any point on a curve is simply the slope of a tangent drawn at that point. The tangent can be drawn easily with a transparent straightedge. The differential, $\frac{dy}{dx}$, is the slope of this tangent.

In figure 26, the slope of the straight line is 5.8 (that is, y increases 5.8 units for each increase of one unit of x). In figure 27, the slope is -0.002 (that is, y decreases 0.002 units for each increase of one unit of x). Thus, in figure 26 $\frac{dy}{dx} = 5.8$, and in figure 27, $\frac{dy}{dx} = -0.002$.

These differentials can be read most easily by drawing the dotted lines shown on the diagrams. These dotted lines are drawn parallel to the tangent and through the origin (the point x=0, y=0). To draw these parallel lines, place one side of a right triangle along the original curve, place a straight-edge along another side of the triangle, and then slip the triangle along the straight edge. With a little practice it is very easy to draw parallel lines.

The slope of the tangent is the same as the slope of the dotted parallel line. It is measured by the height of the dotted line corresponding with one unit on the x-axis. In figure 26 it is 5.8. In figure 27 it would not be possible to read the height of the dotted line corresponding to one unit on the x-axis. So we read the height corresponding to 1,000 units. It is -2. So the slope is -2/1,000 or -0.002.

Graphic differentiation is quick and easy. It is important in any sort of marginal analysis.

We have not given data for these charts as the curves are purely hypothetical and are shown merely to illustrate the method.

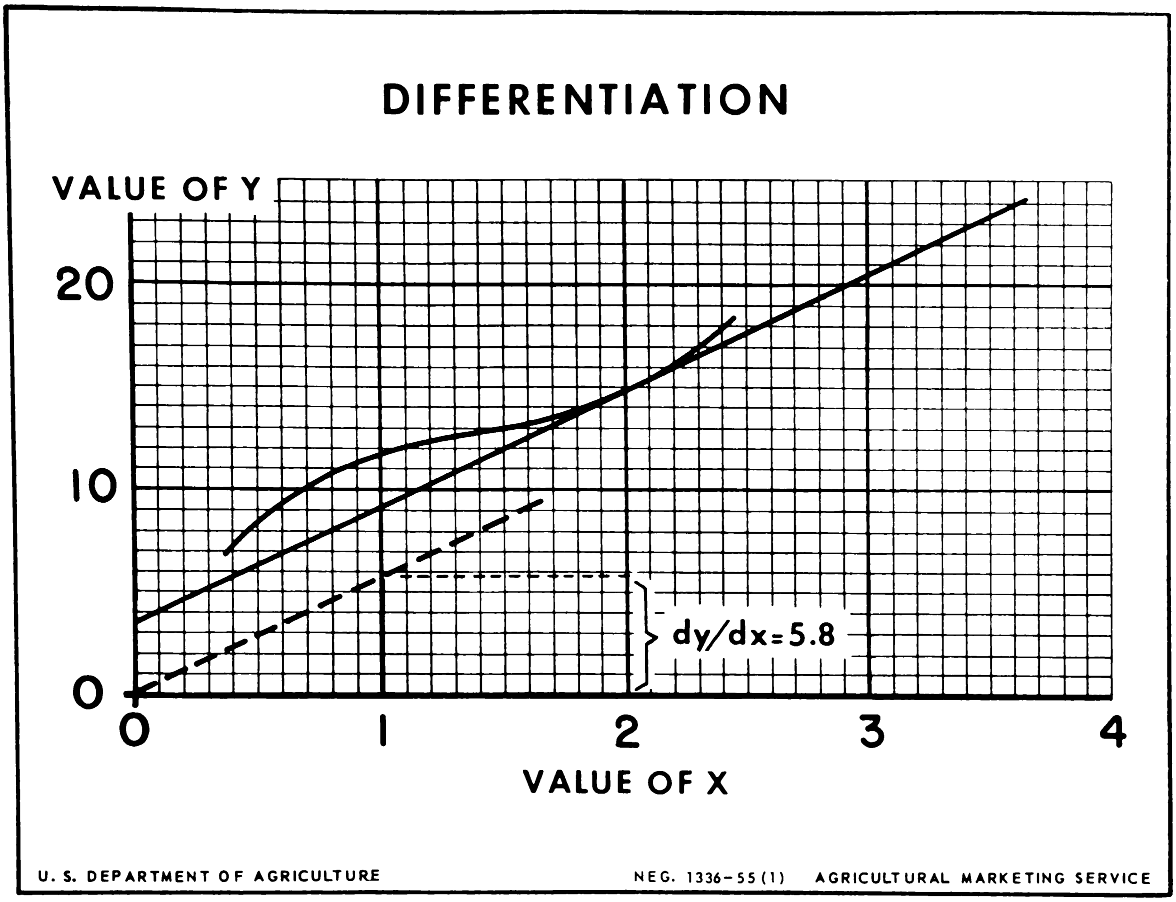


Figure 26

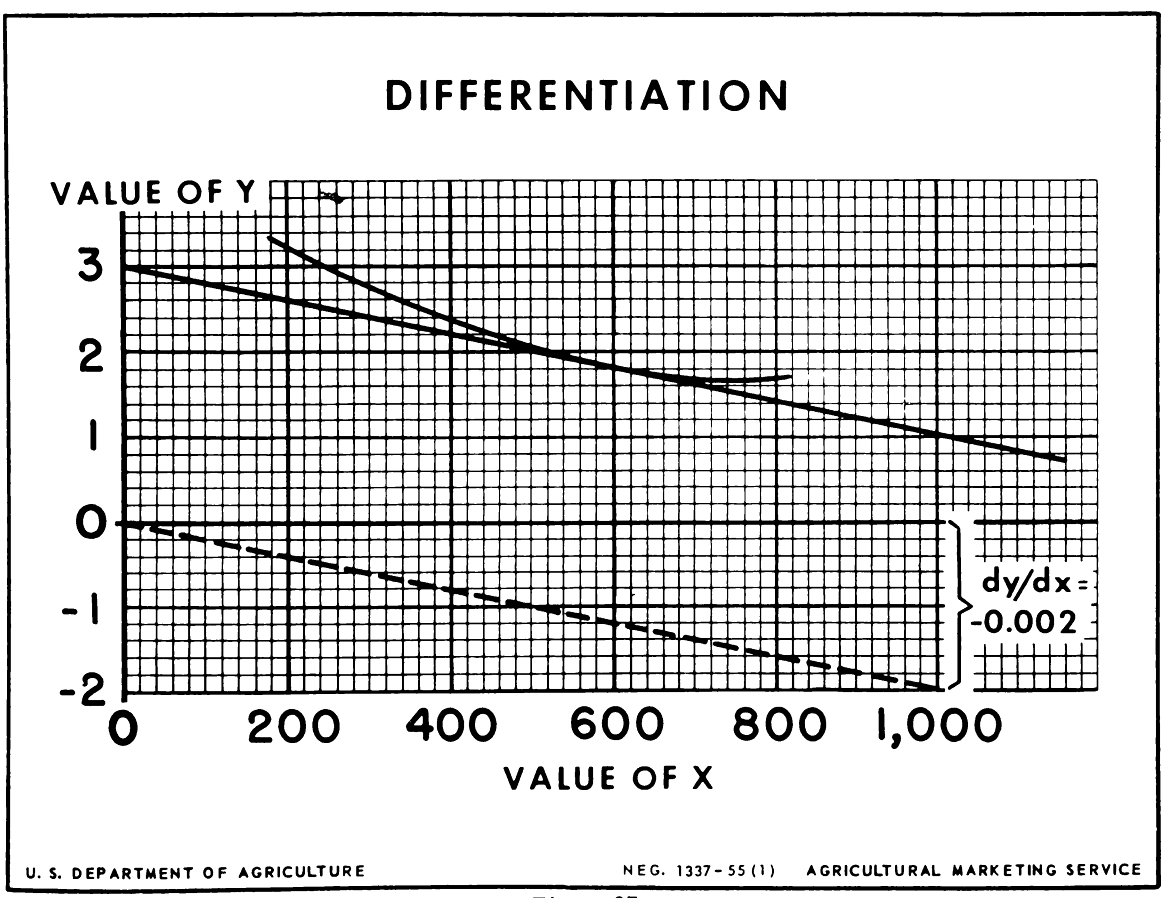


Figure 27

Marginal Returns

Often the economist has a demand curve showing estimates of average prices corresponding with a range of quantities sold. His problem may call for an analysis of marginal returns (or marginal expenditures of consumers). The easiest way to do this is to find graphically several points on the returns curve.

Robinson 17/ explained the geometry of this. Briefly, total returns are $R = pq$. We want $\frac{dR}{dq} = p + \frac{dp}{dq}\, q$. We can take any point on the demand curve, such as point A in our diagram (32 pounds, at an average price of 41 cents), and draw a tangent to the curve at that point. We then draw a line parallel to the tangent such that it cuts the price axis at the price indicated by the point on the demand curve (that is, at 41 cents). This parallel cuts a perpendicular dropped from A at point B, and the price equivalent of B measures the marginal returns corresponding to the quantity sold at point A on the demand curve. Here marginal returns are 10.5 cents when 32 pounds per capita are sold.

This is a simple process and can be done in five seconds. With a little practice you can quickly locate several points on the marginal returns curve, and then draw the whole curve.

17/ Robinson, Joan. The Economics of Imperfect Competition, p. 30. London. 1933.

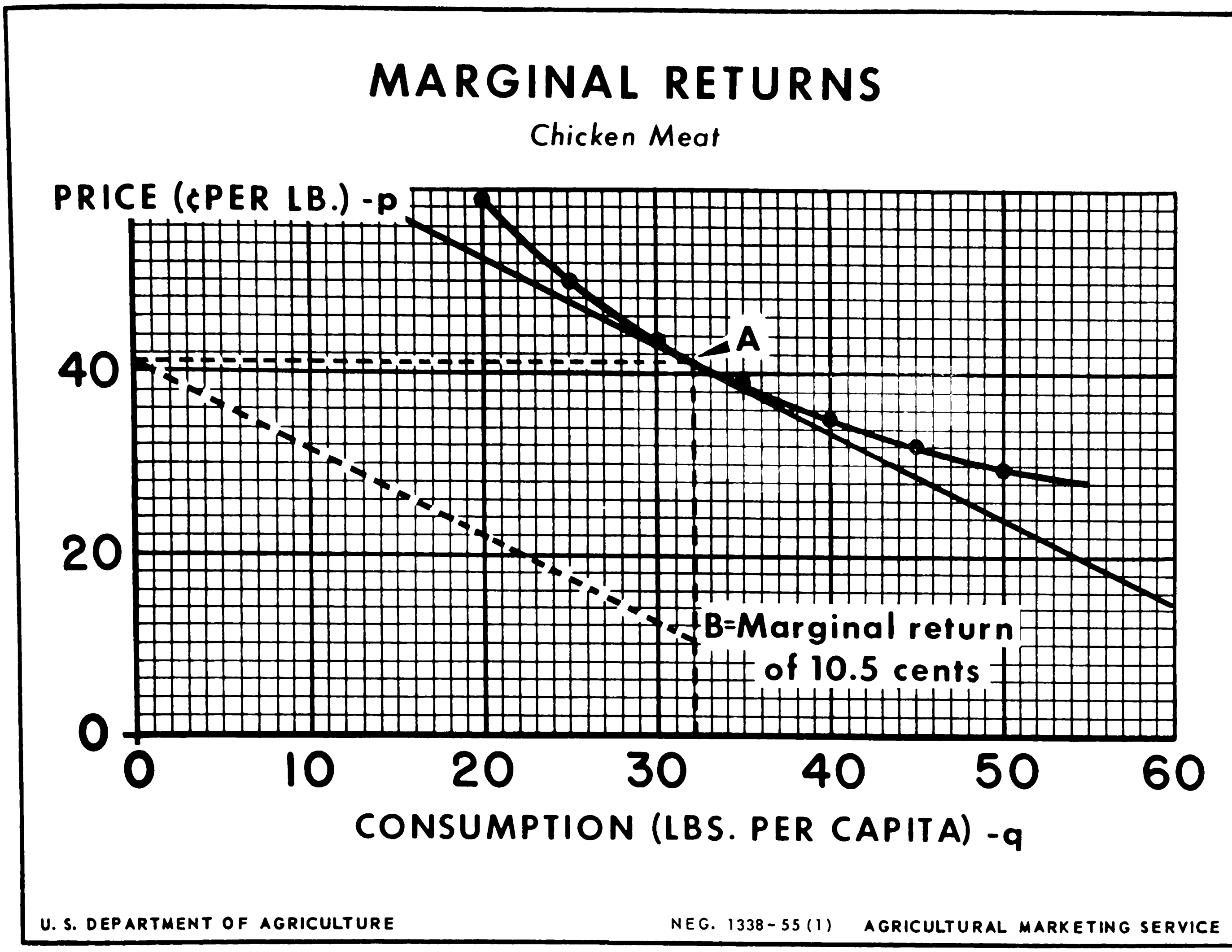


Figure 28

Chicken meat: Price per pound at retail associated with given levels of consumption per capita

Price	Consumption
Cents	Pounds
59.1	20
49.9	25
43.5	30
38.8	35
35.1	40
32.1	45
29.7	50

Regression coefficient based on the reciprocal of an assumed elasticity of demand coefficient of -1.33. See Foote, Richard J. and Fox, Karl A. Analytical Tools for Measuring Demand. U. S. Dept. Agr. Agr. Handbook 64. 1954. p. 40.

Marginal Costs

A marginal cost curve can be obtained from a curve of average costs by the same graphic procedure as that just explained for marginal returns. This process is illustrated in the diagram. In this instance, the marginal curve will be above the average curve. To find the marginal cost at point A in the diagram, we erect a perpendicular line at point A and draw a tangent to the average cost curve at this point. We also draw a horizontal line from point A to the cost axis and note the point at which this line cuts the axis. We then draw a line through this point that is parallel to the tangent. The cost at which this line cuts the perpendicular line is the marginal cost for the input represented by point A.

In the example used here, we show average costs of land and fertilizer per unit of output for given inputs of fertilizer applied to an acre of land. Point A applies to slightly more than $6 worth of fertilizer. For this amount, average costs per unit of output are about $0.237. Marginal costs, as indicated by B, are $0.292. As in the preceding example, several points on the marginal curve can be located as a basis for drawing the entire curve.

If xy is given as a fraction of x, as in these examples, we always can compute

$$\frac{d\,xy}{d\,x} = y + \frac{dy}{dx}\,x$$

by this process no matter what x and y represent.

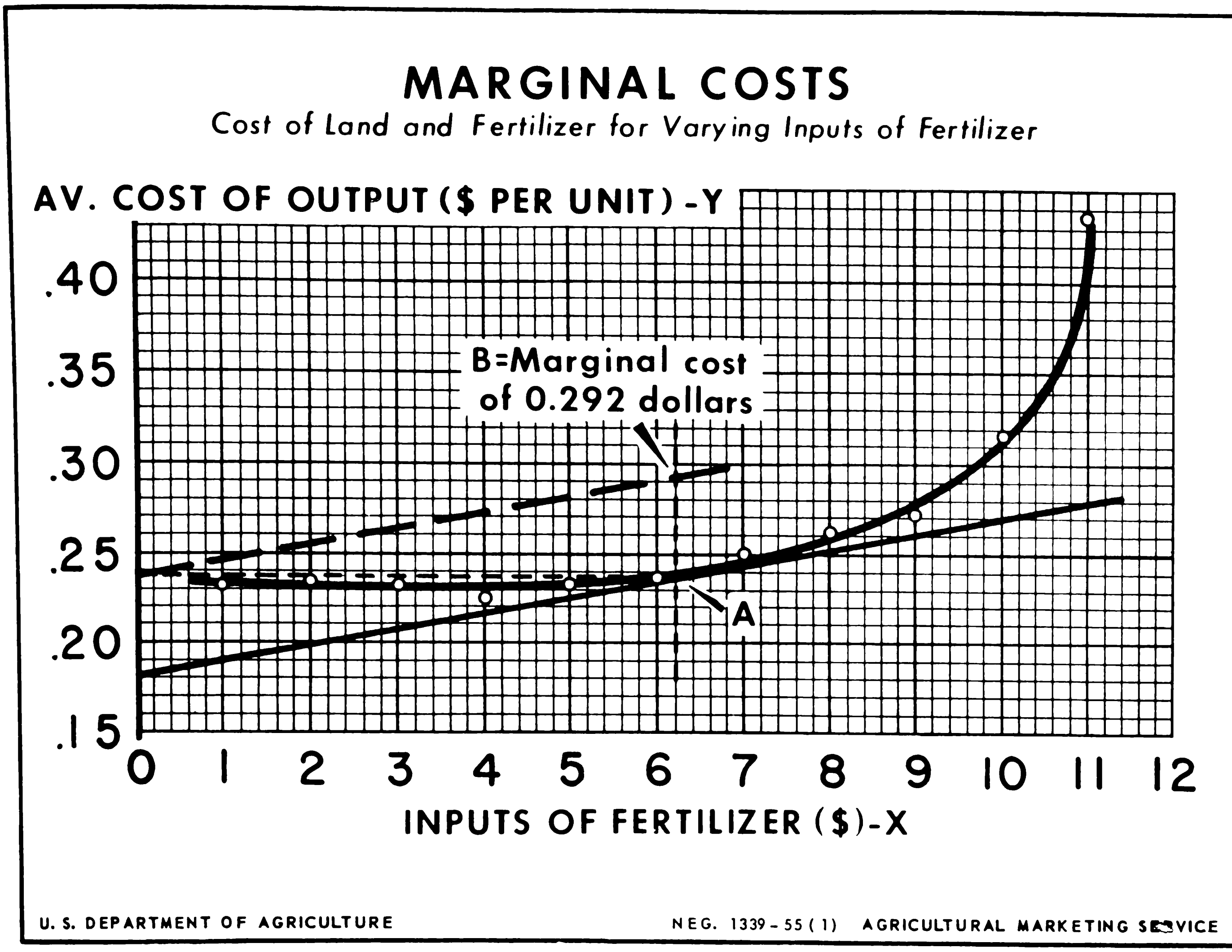


Figure 29

Total output of a given crop and cost per unit of output for given inputs

Cost of total input: Land	Cost of total input: Fertilizer	Total output	Cost per unit of output for - Land	Cost per unit of output for - Fertilizer	Cost per unit of output for - Total
Dollars	Dollars		Dollars	Dollars	Dollars
10	1	47	0.213	0.0213	0.2343
10	2	51	.196	.0392	.2352
10	3	56	.178	.0536	.2316
10	4	62	.161	.0645	.2255
10	5	64	.156	.0781	.2341
10	6	67	.149	.0895	.2385
10	7	68	.147	.1030	.2500
10	8	69	.145	.1159	.2609
10	9	70	.143	.1287	.2717
10	10	64	.161	.1562	.3172
10	11	48	.208	.2294	.4374

Black, John D. Production Economics. New York. 1926. pp. 317-318.

ROOTS OF A POLYNOMIAL

$$x^3 - 1.2240\ x^2 + 0.3695\ x - 0.0183 = 0$$

It may seem strange to discuss the roots of a polynomial in a handbook dealing with graphic analysis. Roots of polynomials are used mainly in high-powered mathematical studies dealing with such things as canonical regression, component analysis, and cyclical variation. But graphics can help, even in these studies.

We have included this diagram to illustrate the use of graphics in connection with more elaborate mathematical techniques. The particular polynomial is taken from Tinter. 18/ Tinter was dealing with a problem of canonical regression. The largest root of the above equation indicates the squared correlation coefficient. We shall not bother to explain how the equation was obtained. We are concerned only with computing its roots--and especially its largest root.

The roots of a polynomial are values of x which satisfy the equation. There are many mathematical tricks for discovering such values of x. But the graphic method illustrated here is practical and easy.

We simply plot several values for x. Thus if $x=0$, the polynomial equals -0.0183; so we plot $y=-0.0183$ corresponding to $x=0$. If $x=0.1$, the polynomial equals 0.0074; so we plot $y=0.0074$ corresponding to $x=0.1$. We proceed to compute several points on the curve, $y=x^3 - 1.2240\ x^2 + 0.3695x - 0.0183$. When we have enough points, we draw a curve through them. Wherever this curve crosses the x-axis, it indicates a real root. In this case, the roots are approximately 0.06, 0.38, and 0.78. The canonical correlation is approximately equal to the square root of 0.78.

We could locate any of these roots more exactly by blowing up the part of the diagram near the root. Thus, we could draw a new diagram for the part of the curve between $x=0.76$ and $x=0.80$, plot the curve on a blown-up scale, and compute the largest root more accurately. This could be repeated until we obtained as many significant figures as wanted.

As a guide to the parts of the curve that must be plotted, we know that there must be as many roots as the degree of the curve. Here we have a third degree polynomial, so we know that there must be three roots. Once we have located them, our job is finished. Sometimes we have multiple roots (that is, two or more roots at a single point) or imaginary roots. These also can be located by graphic means but these topics are beyond the scope of this handbook.

18/ Tinter, Gerhard. Econometrics. New York. 1952. Taken from equations (18) on p. 119, letting $x=\lambda^2$

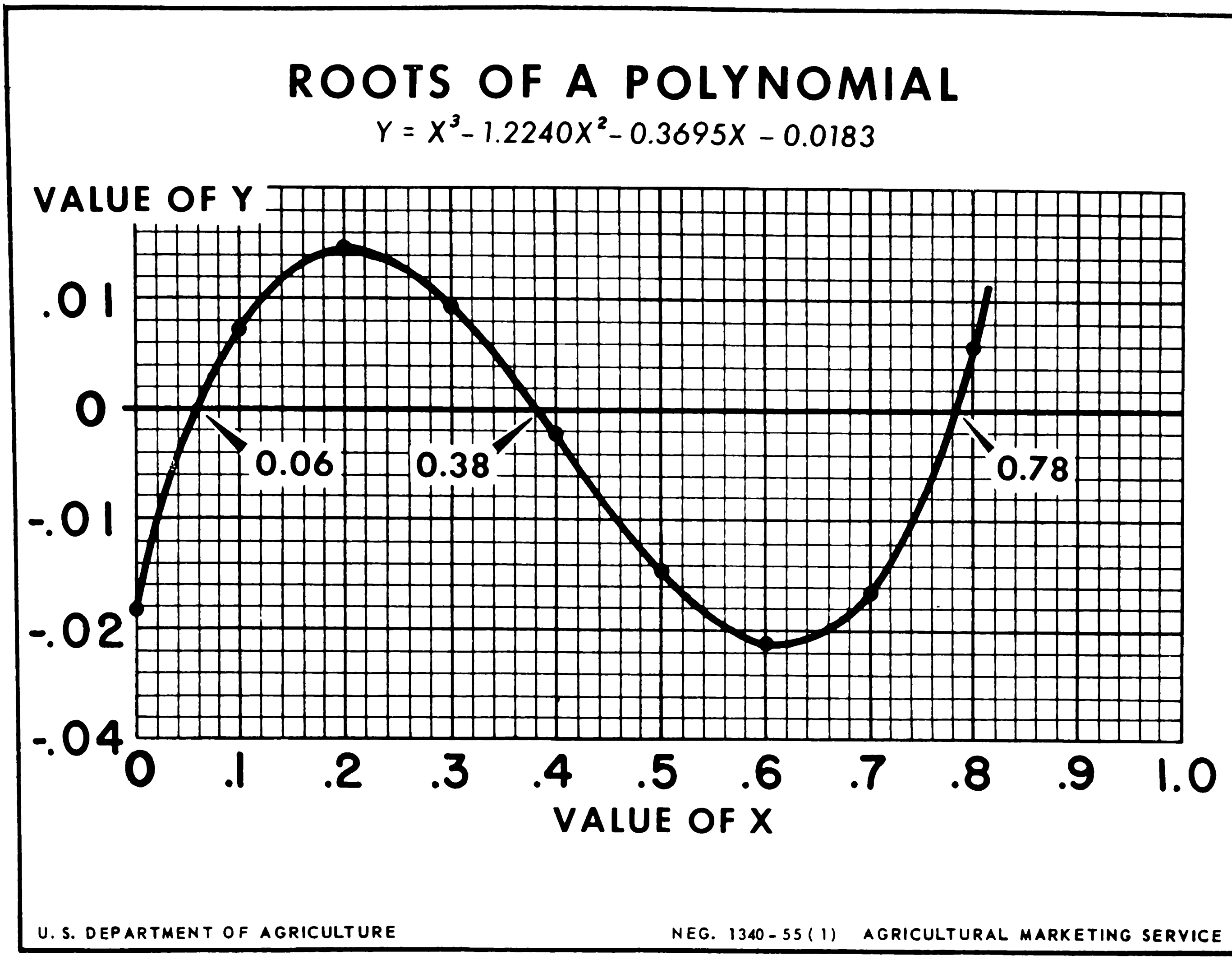


Figure 30

Values of a third-degree polynomial, y, at specified levels of x 1/

x	y
0	-0.0183
.1	.0074
.2	.0146
.3	.0094
.4	- .0023
.5	- .0146
.6	- .0212
.7	- .0164
.8	.0059

1/ $y = x^3 - 1.2240 x^2 + 0.3695 x - 0.0183$.

Data compiled using equations (18) as a basis and letting $x = \lambda^2$. Tintner, Gerhard. Econometrics. New York. 1952. p. 119.

SOLUTION OF SIMULTANEOUS EQUATIONS

Supply and Demand Curves for Winter Tomatoes

We end this handbook with another use of graphics as an aid to mathematical computation. Statisticians often must solve two or more equations simultaneously. Various methods of solution are available, including the popular Gauss-Doolittle technique. But the equations can also be solved graphically. The diagram illustrates only the solution of pairs of equations. It is possible to solve any number of equations graphically by a process very similar to the Gauss-Doolittle method. But we shall not explain the procedure here. 19/

To solve any pair of equations, we substitute several successive values of x in each equation, compute the corresponding values of y, plot the value of y corresponding to each value of x, and draw a smooth curve through the observations. When these operations are performed for each equation, this gives us a pair of curves. Wherever the two curves cross one another, there is a solution of the two equations.

Any pair of linear equations will have one, and only one, real solution--except in the extreme case where the two lines are identical or parallel, where there are infinitely many or no solutions, respectively. Quadratic equations have up to four solutions to a pair of equations, depending on how they are situated one to another. For equations of any degree, solutions are real wherever the curves cross one another; otherwise they are imaginary.

In the linear equations shown here we have a supply curve and a demand curve for winter tomatoes. 20/ The scales in this chart refer to logarithms. The supply curve shows the quantity of tomatoes that will be imported with given prices. It is quite steep, indicating that large changes in prices are required to have much effect on imports. The demand curve shows the relation between domestic prices and the quantity imported. It is highly elastic, reflecting the high degree of competition between domestic and imported tomatoes. The values in 1952 of the other factors that affect imports of tomatoes and their prices have been combined with the constant values in the respective equations to give these two equations that show directly the simultaneous relations between imports and price. Since none of the coefficients in these equations differs significantly from zero, little confidence should be placed in their economic meaning. However, they serve as a good example of the simultaneous solution of a pair of equations.

Linear relations are shown for this example because these are common in economic analysis. Graphic solutions, however, are much more useful for approximating simultaneous solutions for more complicated equations.

19/ The method is described in Maxfield, John E. and Waugh, Frederick V. A Graphic Solution to Simultaneous Linear Equations. Math. Tables and Other Aids to Computations, 5:246-248, illus. 1951.

20/ This analysis is taken from Shuffett, D. Milton. The Demand and Price Struction for Selected Vegetables. U. S. Dept. Agr. Tech. Bul. 1105, pp. 107-108. 1954.

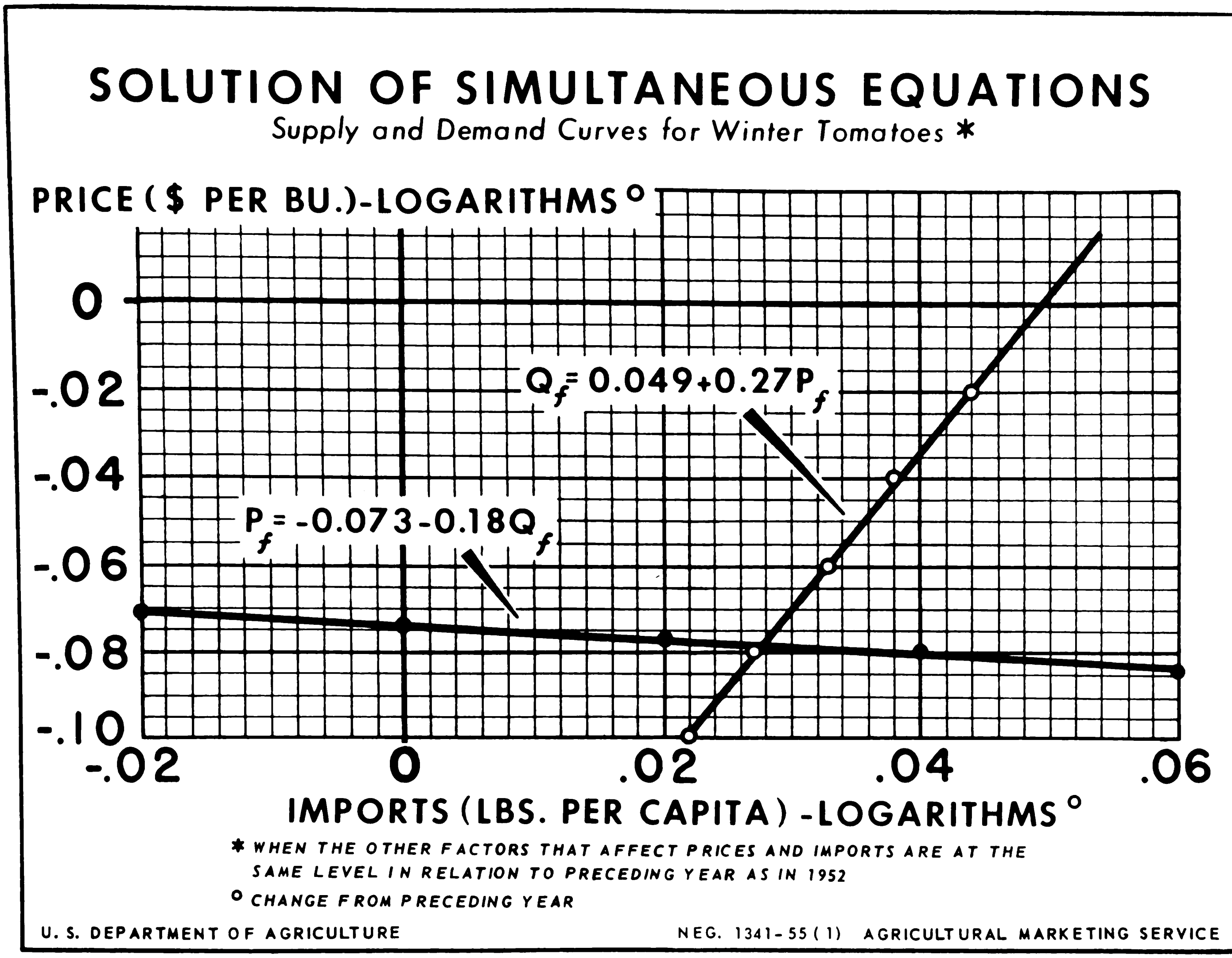


Figure 31

Winter tomatoes: Associated prices received by farmers and quantities imported, change from preceding year

Demand equation		Supply equation	
Price 1/	Quantity	Quantity 1/	Price
Logarithm	Logarithm	Logarithm	Logarithm
-0.071	-0.02	0.022	-0.10
- .073	0	.027	- .08
- .077	.02	.033	- .06
- .080	.04	.038	- .04
- .084	.06	.044	- .02

1/ When other variables that affect prices or imports, respectively, are at the same level in relation to the preceding year as in 1952.

Shuffett, D. Milton. The Demand and Price Structure for Selected Vegetables. U. S. Dept. Agr. Tech. Bul. 1105. 1954. pp. 107 and 108.